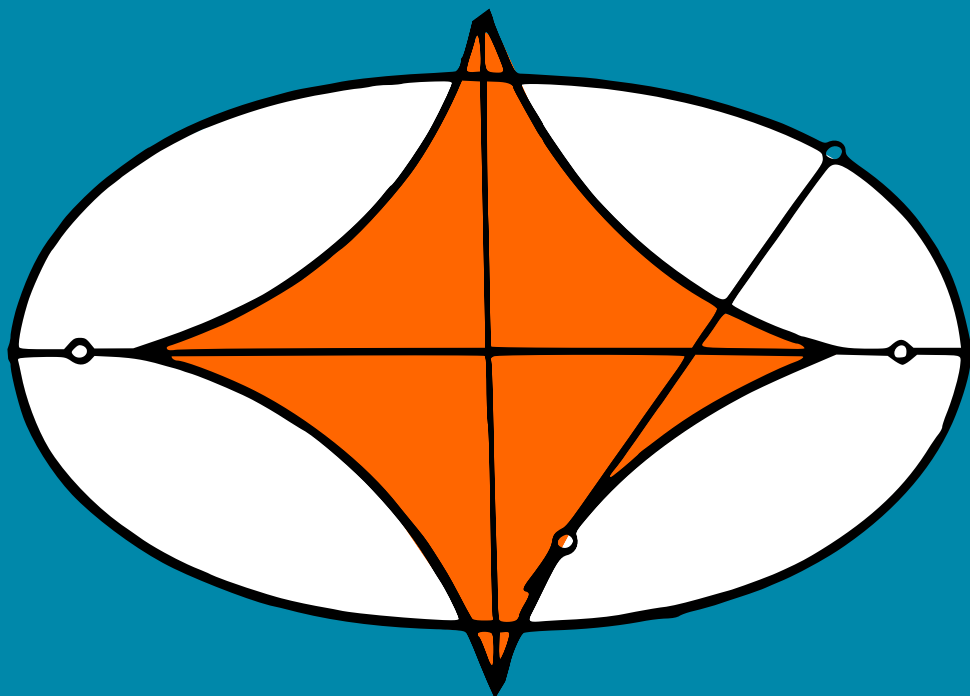
Y. Bougrov, S. Nikolski

# COURS DE MATHÉMATIQUES SUPÉRIEURES

Tome 1



Éditions Mir Moscou

Я. С. БУГРОВ, С. М. НИКОЛЬСКИЙ

# ВЫСШАЯ МАТЕМАТИКА

## ДИФФЕРЕНЦИАЛЬНОЕ И ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ

## ЭЛЕМЕНТЫ ЛИНЕЙНОЙ АЛГЕБРЫ И АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

Y. BOUGROV, S. NIKOLSKI

# COURS DE MATHÉMATIQUES SUPÉRIEURES

## TOME I

### CALCUL DIFFÉRENTIEL ET INTÉGRAL

### ÉLÉMENTS D'ALGÈBRE LINÉAIRE ET DE GÉOMÉTRIE ANALYTIQUE

ÉDITIONS DE MOSCOU

Traduit du russe par
Djilali EMBAREK

*На французском языке*

# TABLE DES MATIÈRES

# AVANT-PROPOS

Le cours de mathématiques que nous proposons est composé de deux tomes.

Dans le premier tome, les trois premiers chapitres sont consacrés aux notions élémentaires d'analyse, quatre autres chapitres au calcul différentiel et intégral des fonctions d'une variable, un chapitre au calcul différentiel des fonctions de plusieurs variables, un chapitre aux séries et un chapitre enfin à l'algèbre linéaire (théorie des déterminants et des matrices, systèmes linéaires d'équations, opérateurs linéaires, opérateurs hermitiens, formes quadratiques) et à la géométrie analytique dans le plan et dans l'espace (droite, plan, coniques, quadriques).

Les auteurs ont conscience que le programme de mathématiques du second cycle est surchargé et qu'il n'est point besoin par conséquent de tout prouver. Ils espèrent que le professeur aidera les élèves à procéder aux coupes nécessaires.

Plusieurs paragraphes du chapitre premier traitent du nombre réel. Nous pensons que l'approche du nombre réel comme une forme décimale est intéressante. La lecture de la démonstration du lemme 2 relatif à une suite de formes décimales bornée croissante est souhaitable. Cependant dans l'étude de ces questions on peut se limiter au § 7 où sont énumérées les propriétés du nombre réel et au § 8 qui traite de l'approche axiomatique de la notion de nombre.

Les propositions sont en règle générale démontrées. Cependant dans les cas où les démonstrations sont omises en dimension $n$, des explications détaillées indiquent la marche à suivre en dimension deux et trois.

Les auteurs attirent l'attention sur le fait que les neuf premiers chapitres et le chapitre 10 « Eléments d'algèbre linéaire et de géométrie analytique » s'imbriquent étroitement et leur lecture doit être menée de pair. D'autre part avant d'aborder le chapitre 8 sur les fonctions de plusieurs variables, il est vivement conseillé de s'initier aux notions de base de l'espace à $n$ dimensions.

Par ailleurs l'étude de l'opérateur hermitien et des formes quadratiques passe par celle des propriétés des fonctions continues sur un ensemble fermé (§ 12, chapitre 8).

Le paragraphe consacré à l'extrémum des fonctions de plusieurs variables implique la connaissance des formes quadratiques.

Les auteurs espèrent que cet ouvrage aidera les élèves des écoles techniques à parfaire leurs connaissances mathématiques.

CHAPITRE PREMIER

# INTRODUCTION

## § 1. Objet des mathématiques. Grandeurs variables et constantes, ensembles

Les mathématiques occupent une place particulière dans le concert des sciences. Les mathématiques se définissent comme la science des formes de l'espace et des rapports quantitatifs du monde réel. Certes, si l'on tient compte de l'état actuel des mathématiques et de la diversité de ses structures, il faut comprendre ces formes et ces rapports dans leur acception la plus large.

Les mathématiques arment les autres sciences du langage des nombres et des symboles pour exprimer toute sorte de rapports entre les phénomènes de la nature. Mais avant de faire usage des mathématiques le biologiste, le physicien ou l'économiste doit bien concevoir le phénomène étudié et le décomposer en parties susceptibles d'être traitées par les mathématiques.

Les mathématiques étudient les modèles logiques conçus pour décrire les phénomènes naturels et sociaux ainsi que les rapports existant entre les éléments de ces modèles. Si le modèle mathématique décrit correctement le phénomène étudié, il permet de dévoiler les lois qui régissent ce phénomène, autrement dit les mathématiques sont capables de mettre à jour les aspects qualitatifs du phénomène étudié.

Vu son niveau élevé d'abstraction, un modèle mathématique est capable de simuler différents processus. Ainsi une même équation différentielle décrit la désintégration radioactive et les variations de température d'un corps.

Quand on étudie des phénomènes naturels et sociaux, on se heurte constamment à des variations de grandeurs, à des relations entre grandeurs. Donc, la notion de grandeur variable est fondamentale en analyse mathématique.

Par *grandeur variable* on entendra une grandeur qui dans le processus étudié est susceptible de prendre au moins deux valeurs distinctes. La grandeur qui ne prend qu'une valeur sera dite *constante*.

F. Engels a signalé que l'emploi de la variable cartésienne a introduit le mouvement et la dialectique en mathématiques.

Si l'on regroupe toutes les valeurs prises par une variable, on obtient l'*ensemble* des valeurs de cette variable.

La notion d'*ensemble* est aussi fondamentale en mathématiques. C'est une notion liminaire que nous n'essayerons même pas de définir par d'autres notions simples.

Par *ensemble* on entendra une collection d'objets de nature quelconque.

On peut parler de l'ensemble des étudiants d'une faculté, de l'ensemble des molécules d'un corps, de l'ensemble des T.V. couleurs dans une salle, etc. Les objets d'un ensemble s'appellent *éléments* de cet ensemble.

On désignera les ensembles par des lettres majuscules $A, B, \ldots$ $\ldots, X, Y, \ldots$ et leurs éléments par des lettres minuscules $a, b, \ldots$ $\ldots, x, y, \ldots$

Si un élément $x$ appartient à un ensemble $A$, on note ceci par: $x \in A$. Si $x$ n'appartient pas à $A$, on écrit: $x \notin A$. Le symbole $A \subset B$ (lire $A$ compris dans $B$) signifie que tout élément de $A$ est élément de $B$.

L'ensemble $A$ s'appelle alors sous-ensemble de $B$. On se sert encore de la notation équivalente: $B \supset A$ (lire l'ensemble $B$ comprend l'ensemble $A$). Les symboles $\subset$ et $\supset$ sont les *signes d'inclusion*.

Si un ensemble ne contient aucun élément, on dit qu'il est *vide* et on le note $\varnothing$. Il est évident que $\varnothing \subset A$, où $A$ est un ensemble quelconque.

Pour noter les ensembles on se sert d'accolades à l'intérieur desquelles on inscrit par des procédés divers les éléments qui les composent. L'expression $\mathbf{N} = \{1, 2, 3 \ldots\}$ désigne l'*ensemble des entiers naturels*; $\{0, 1, 2, \ldots\}$, l'ensemble des entiers positifs; $\mathbf{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, l'*ensemble des entiers relatifs*. Par exemple, l'ensemble $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 0\}$ est composé des chiffres du système décimal. Il est évident que $2 \in A$ et $\frac{1}{2} \notin A$.

On dit que deux ensembles $A$ et $B$ sont *égaux*, et on note $A = B$, si $A \subset B$ et $B \subset A$. Il n'est pas toujours facile d'établir l'égalité de deux ensembles. Par exemple, si $A = \{6, 8, 10, \ldots\}$ et $B = \{p + q\}$ est l'ensemble des sommes des nombres premiers $p$ et $q$ plus grands que 2, il est clair que $B \subset A$. Cependant, à ce jour on ne sait pas encore si $A \subset B$, c'est-à-dire si tout nombre entier pair $\geqslant 6$ se représente par une somme de nombres premiers $> 2$.

Dans cet ouvrage on aura principalement affaire aux ensembles numériques.

## § 2. Opérations sur les ensembles

Les ensembles sont justiciables des opérations arithmétiques d'addition et de multiplication qui jouissent de propriétés en beaucoup de points similaires à celles de l'addition et de la multiplication des nombres.

Soient donnés deux ensembles arbitraires $A$ et $B$. On appelle *somme* ou *réunion* des ensembles $A$ et $B$ l'ensemble $C$ composé des éléments des ensembles $A$ et $B$, et on note: $C = A + B$ ou $C = A \cup B$ (fig. 1). Il est immédiat que $A + A = A$.

On appelle *produit* ou *intersection* des ensembles $A$ et $B$ l'ensemble $C$ composé des éléments appartenant à la fois à $A$ et à $B$,

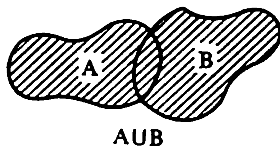



Fig. 1

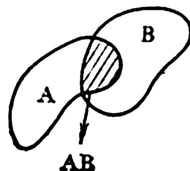



Fig. 2

et on note $C = AB$ ou $C = A \cap B$ (fig. 2). Il est évident que $A \cap A = A$. Si $AB = \varnothing$, on dit que les ensembles $A$ et $B$ sont *disjoints*. En se servant de la notion d'égalité des ensembles on démontre que:

1) $A + B = B + A$;
2) $(A + B)\, C = AC + BC$;
3) $(AB)\, C = A\, (BC)$;
4) $(A + B) + C = A + (B + C)$.

Prouvons 2). Si $x \in (A + B)\, C$, c'est que, par définition du produit, $x \in A + B$ et $x \in C$. De la définition de la somme il s'ensuit que $x \in A$ ou $x \in B$. Supposons pour fixer les idées que $x \in A$. Alors $x \in AC$ et par suite $x \in AC + BC$. Donc, $(A + B)\, C \subset AC + BC$. Si à présent $x \in AC + BC$, alors ou bien $x \in AC$. ou bien $x \in BC$. Supposons que $x \in AC$. Alors $x \in A + B$ et $x \in C$, c'est-à-dire que $x \in (A + B)\, C$. D'où $AC + BC \subset (A + B)\, C$. Ce qui prouve 2).

On appelle *différence* des ensembles $A$ et $B$ l'ensemble $R = A \setminus B$ composé des éléments de $A$ n'appartenant pas à $B$.

On remarquera que dans le cas général $(A \setminus B) + B \neq A$ (fig. 3). Mais si $B \subset A$, alors $(A \setminus B) + B = A$ (fig. 4).

Les ensembles et les opérations d'addition et de multiplication que nous venons de voir forment une algèbre originale privée de coefficients et de puissances.

## § 3. Symbolique de la logique mathématique

Pour abréger les notations on se servira dans la suite de quelques symboles élémentaires de la logique. Si une proposition nous intéresse non pas par son contenu mais par ses relations avec d'autres, on la désignera par $\alpha$, $\beta$, . . . . La notation $\alpha \Longrightarrow \beta$ signifie que la proposition $\alpha$ entraîne la proposition $\beta$, la notation $\alpha \Longleftrightarrow \beta$, que les propositions $\alpha$ et $\beta$ sont équivalentes, c'est-à-dire que $\alpha$ entraîne $\beta$ et $\beta$ entraîne $\alpha$.

La notation $\forall x \in A : \alpha$ signifie que la proposition $\alpha$ a lieu pour tout $x \in A$. Le symbole $\forall$ s'appelle *quantificateur universel*.

La notation $\exists\, y \in B : \beta$ signifie qu'il existe un élément $y \in B$ pour lequel la proposition $\beta$ a lieu. Le symbole $\exists$ s'appelle *quantificateur existentiel.*

Le symbole $\overline{\alpha}$ désigne la *négation* de la proposition $\alpha$ ou plus brièvement la proposition « non $\alpha$ ».

Construisons la négation de la proposition $\forall\, x \in A : \alpha$.

Si cette proposition n'a pas lieu, la proposition $\alpha$ n'est pas réalisée pour tous les $x \in A$, en d'autres termes il existe un élément $x \in A$ pour lequel $\alpha$

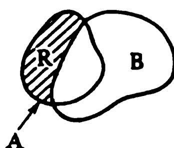
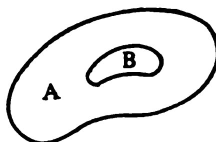


Fig. 3



$(A \setminus B) + B = A$

Fig. 4

n'a pas lieu : $\overline{\forall\, x \in A : \alpha} \Leftrightarrow \exists x \in A : \overline{\alpha}$. De façon analogue $\overline{\exists\, y \in B : \beta} \Leftrightarrow \forall y \in$ $\in B : \overline{\beta}$.

Donc, pour construire la négation d'une formule contenant les symboles $\forall$ et $\exists$, il faut remplacer $\forall$ par $\exists$ et $\exists$ par $\forall$ et surmonter de la barre de négation la propriété qui suit les deux points. Ainsi la négation de la proposition

$$\exists M,\ \forall x \in A : f(x) \leqslant M$$

est

$$\overline{\exists M,\ \forall x \in A : f(x) \leqslant M} \Leftrightarrow \forall M,\ \exists x \in A : \overline{f(x) \leqslant M} \Leftrightarrow$$
$$\Leftrightarrow \forall M,\ \exists x \in A : f(x) > M.$$

## § 4. Nombres réels

La notion de nombre est liminaire et fondamentale en mathématiques. Au cours des siècles, cette notion a pris des extensions successives. L'ensemble des entiers naturels

$$\mathbf{N} = \{1, 2, 3, \ldots, n, \ldots\}$$

est apparu en raison du comptage des objets. Les besoins de la pratique et du développement des mathématiques ont poussé les mathématiciens à introduire les entiers relatifs

$$\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

et les nombres rationnels

$$\mathbf{Q} = \{m/n\}, \quad \text{où} \quad m, n \in \mathbf{Z}, \quad n \neq 0.$$

Pour représenter un nombre rationnel de manière unique, on admettra, sauf mention expresse du contraire, que la fraction $m/n$ est irréductible.

L'introduction des nombres rationnels n'a toutefois pas entièrement résolu un important problème de mesure des segments. Il

existe en effet des segments dont la longueur n'est pas rationnelle. Exemple: la diagonale d'un carré de longueur 1.

D'où la nécessité de considérer d'autres nombres: les *nombres irrationnels*. L'ensemble des nombres rationnels et irrationnels constitue l'ensemble **R** des *réels*. Il existe plusieurs procédés d'introduction (de définition) des réels. Nous optons pour celui qui consiste à les représenter par des formes décimales illimitées

$$a = \pm a_0,\ a_1a_2a_3 \ldots, \tag{1}$$

$a_0$ étant un entier non négatif. $a_k$ des décimales. Donc, $a_k$ ne peut prendre que l'une des valeurs 0, 1, 2, . . ., 9. On omet souvent le signe $+$.

Le nombre 0 s'écrit sous l'une des formes suivantes:

$$0 = +0{,}00 \ldots = 0{,}00 \ldots = -0{,}00 \ldots$$

Pour représenter un nombre rationnel non nul $\pm m/n$ ($m > 0$, $n > 0$) par une forme décimale, on divise tout simplement $m$ par $n$ comme à l'école primaire:

$$\begin{array}{r|l} m & n \\ \cline{2-2} & a_0,\ a_1a_2 \ldots \end{array} \tag{2}$$

On remarquera que si l'on applique ce procédé à une autre représentation de la fraction $\pm mp/np = \pm m/n$ ($p > 0$), on obtient le même résultat.

On pose

$$\pm \frac{m}{n} = \pm a_0,\ a_1a_2a_3 \ldots \tag{3}$$

et on appelle le second membre de (3) *symbole* (ou *forme* ou *développement*) *décimal du nombre* $\pm m/n$.

Si le dénominateur de la fraction est de la forme $n = 2^s5^l$, où $s$ et $l$ sont des entiers positifs, alors l'opération (2) s'achève au bout d'un nombre fini de pas et on obtient un *symbole décimal limité*

$$\pm \frac{m}{n} = \pm a_0,\ a_1 \ldots a_M = \pm a_0,\ a_1 \ldots a_M00 \ldots \quad (a_M > 0). \tag{4}$$

On utilise encore une autre représentation de la fraction (4):

$$\pm a_0,\ a_1 \ldots a_M = \pm a_0,\ a_1 \ldots a_{M-1}(a_M - 1)\,99 \ldots =$$
$$= \pm a_0,\ a_1 \ldots a_{M-1}(a_M - 1)\,(9), \tag{5}$$

bien qu'elle ne provienne pas du processus (2). Le troisième membre de (5) est un *symbole décimal illimité*.

On dira qu'un symbole décimal

$$\beta_0,\ \beta_1\beta_2 \ldots$$

est *illimité* si pour tout entier naturel $n$ on peut exhiber un entier naturel $k > n$ tel que le chiffre décimal $\beta_k$ soit $> 0$.

Ainsi donc, tout symbole décimal limité peut être, grâce à (5), représenté par un symbole décimal illimité. Les formes décimales dont il sera question plus bas engendreront d'autres nombres.

Supposons maintenant que le dénominateur de la fraction étudiée n'est pas de la forme $2^s5^l$. Alors l'opération (2) est illimitée : à chaque pas on obtient un reste positif. Chaque reste est inférieur à $n$, donc parmi les $n$ premiers restes il en existe deux au moins qui sont égaux. A partir de là l'opération se reproduit et devient *périodique*. Donc, la représentation décimale d'un nombre rationnel est de la forme

$$\pm \frac{m}{n} = \pm a_0, a_1 \ldots a_M b_1 \ldots b_s b_1 \ldots b_s \ldots =$$
$$= \pm a_0, a_1 \ldots a_M (b_1 \ldots b_s) \quad (s < n). \tag{6}$$

Le nombre $b_1 \ldots b_s$ s'appelle *période*. La représentation (5) peut être considérée comme un cas particulier de (6).

EXEMPLES.

$$\left\{\begin{aligned} &\frac{1}{6} = 0{,}166 \ldots = 0{,}1\,(6), \quad \frac{1}{7} = 0{,}(142857), \\ &\frac{2}{9} = 0{,}22 \ldots = 0{,}(2), \\ &\frac{7}{99} = 0{,}0707 \ldots = 0{,}(07), \\ &\frac{7}{990} = 0{,}00707 \ldots = 0{,}0(07). \end{aligned}\right. \tag{7}$$

La représentation (6) s'appelle *forme* (ou *symbole* ou *développement*) *décimale illimitée périodique*.

Ainsi, tout nombre rationnel non nul peut être représenté grâce à l'opération (2), ou grâce à (5) dans le cas d'un nombre décimal, par une forme décimale illimitée périodique. On démontre qu'à tout nombre rationnel correspond une forme décimale illimitée périodique et inversement que toute forme décimale illimitée périodique (6) est engendrée grâce à (2) et à (5) par un nombre rationnel donné par la formule :

$$\pm \frac{m}{n} = \pm \left[ a_0, a_1 \ldots a_M + \frac{\beta_1 \ldots \beta_s}{9 \ldots 9} 10^{-M} \right].$$

Par exemple

$$1{,}237\,(06) = 1{,}237 + 0{,}000\,(06) = 1{,}237 + \frac{06}{99} 10^{-3} = 1{,}237 + \frac{6}{99\,000}.$$

Il existe des formes décimales qui ne sont pas périodiques. Exemple : 0,1010010001 . . . ; 0,121122111222 . . .

Autre exemple : si l'on extrait la racine carrée de 2 d'après la règle habituelle, on obtient la forme décimale illimitée non périodique : $\sqrt{2} = 1{,}41\ldots$ Cette forme est définie en ce sens qu'à tout entier naturel $k$ correspond un chiffre $\alpha_k$ occupant le rang $k$ à droite de la virgule et se calculant de façon unique d'après la règle d'extraction de la racine carrée.

L'analyse mathématique nous offre plusieurs procédés de calcul du nombre $\pi$ avec n'importe quelle précision. On obtient pour le nombre $\pi$ une forme décimale illimitée bien déterminée qui n'est pas une forme décimale périodique mixte.

Donnons maintenant une définition purement formelle d'un nombre irrationnel. *On appelle nombre irrationnel un nombre déterminé par une forme illimitée non périodique*

$$a = \pm\alpha_0,\ \alpha_1\alpha_2\alpha_3\ldots, \tag{8}$$

*où* $\alpha_0$ *est un entier positif,* $\alpha_k$ $(k = 1, 2, \ldots)$ *des chiffres, le signe* « = » *veut dire que nous avons désigné le second membre de* (8) *par* $a$.

*Les nombres rationnels et irrationnels sont dits réels.*

De ce qui précède il suit que *tout nombre réel non nul peut être représenté par une forme décimale illimitée* (8). *Si cette forme est illimitée périodique, ce nombre est rationnel ; si elle est illimitée non périodique, il est irrationnel.*

Une *fraction décimale non nulle* peut être représentée par une forme limitée, mais elle ne définit pas un nouveau nombre rationnel : en effet d'après la convention exprimée par l'égalité (5) elle peut être remplacée par une forme illimitée périodique.

Un *nombre* $a$ *dont les* $\alpha_k$ *ne sont pas tous nuls est positif ou négatif* selon qu'on aura le signe « + » ou le signe « — » dans (8) ; ceci étant on omettra comme toujours le signe « + ».

Les nombres réels ont été définis jusqu'ici d'une manière formelle. Il reste encore à définir les opérations arithmétiques sur eux, à introduire la notion « $>$ » et à vérifier que ces opérations et cette notion sont compatibles avec les opérations et la notion « $>$ » respectives, relatives aux nombres rationnels. De plus il faut s'assurer qu'ils sont doués des mêmes propriétés que les nombres rationnels.

## § 5. Définition de l'égalité et de l'inégalité

Soient deux nombres $a = \pm\alpha_0,\ \alpha_1\alpha_2\ldots$, $b = \pm\beta_0,\ \beta_1\beta_2\ldots$ définis par des formes décimales *illimitées*. On dira que ces nombres sont *égaux* si et seulement s'ils ont même signe et

$$\alpha_k = \beta_k \quad (k = 0, 1, 2, \ldots).$$

Soient $a$ et $b$ des nombres strictement positifs. Par définition, $a < b$ ou, ce qui revient au même, $b > a$, si $\alpha_0 < \beta_0$ ou s'il existe un indice $l$ tel que $\alpha_k = \beta_k$ $(k = 0, 1, \ldots, l)$ et $\alpha_{l+1} < \beta_{l+1}$.

Si les formes décimales illimitées représentant les nombres $a$ et $b$ ont pour expression

$$\begin{cases} a=\alpha_0,\ \alpha_1 \ldots \alpha_N 99 \ldots & (\alpha_N<9), \\ b=\beta_0,\ \beta_1 \ldots \beta_{N_1} 99 \ldots & (\beta_{N_1}<9), \end{cases} \tag{1}$$

on peut représenter ces nombres par les formes décimales limitées

$$\begin{cases} a=\alpha_0,\ \alpha_1 \ldots \alpha_{N-1}(\alpha_N+1), \\ b=\beta_0,\ \beta_1 \ldots \beta_{N_1-1}(\beta_{N_1}+1). \end{cases} \tag{1'}$$

Il est immédiat de voir que si $a = b$ ou $a < b$ au sens de la définition précédente (dans le langage des formes décimales illimitées (1)), alors $a = b$ ou respectivement $a < b$ au sens de l'égalité ($a = b$) ou de l'inégalité ($a < b$) des formes décimales (1').

Par définition, $a > 0$ ou $a < 0$ selon que $a$ est strictement positif ou négatif; par ailleurs, par définition, $a < b$ si $a < 0$ et $b > 0$ ou si $a$, $b < 0$ et $|a| > |b|$.

Si $a = \pm\alpha_0,\ \alpha_1\alpha_2 \ldots$, alors par définition $-a = \mp\alpha_0,\ \alpha_1\alpha_2 \ldots$ et la valeur absolue $|a| = +\alpha_0,\ \alpha_1\alpha_2 \ldots = \alpha_0,\ \alpha_1\alpha_2 \ldots$ Donc,

$$|-a|=|a|=\begin{cases} a & (a\geqslant 0), \\ -a & (a\leqslant 0). \end{cases}$$

On sait qu'entre les nombres réels et les points d'une droite on peut établir une correspondance biunivoque ($\leftrightarrow$) d'après la règle

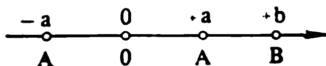



Fig. 5

suivante. Au nombre 0 on associe un point quelconque $O$ sur la droite appelé point zéro et réciproquement. Pour unité de mesure on prend la longueur d'un intervalle quelconque. A chaque nombre réel $\pm a$ ($a > 0$) on associe le point de la droite situé à une distance $a$ à droite de $O$ pour $+a$, et à gauche de $O$, pour $-a$ (fig. 5). Réciproquement, si $A$ est un point arbitraire de la droite situé à une distance $a$ à droite de $O$, on admet qu'il correspond au nombre réel $+a$. S'il est situé à gauche de $O$, on admet qu'il est associé au nombre $-a$. La droite considérée s'appelle *droite numérique* ou axe des réels. Dans la suite on ne fera pas de distinction entre les points de la droite numérique et les nombres réels correspondants. Signalons que la distance de deux points $a$ et $b$ est égale à $|a - b|$ (la définition de la différence est donnée au § 6).

## § 6. Définition des opérations arithmétiques

Supposons qu'à tout entier positif (un *indice*) est associé un nombre $x_n$ d'après une loi. La collection

$$x_0, x_1, x_2, \ldots \tag{1}$$

s'appelle *suite* de nombres. Les nombres $x_n$ sont les *éléments* ou les *termes* de la suite (1). Les éléments $x_n$ et $x_m$ ($m \neq n$) sont distincts, même si éventuellement ils sont égaux.

On dit qu'une *suite* est *croissante* (resp. *décroissante*) si $x_k \leqslant x_{k+1}$ (resp. $x_k \geqslant x_{k+1}$) pour tout $k = 0, 1, 2, \ldots$

On dit qu'une *suite* est *majorée* (resp. *minorée*) s'il existe un nombre $M$ (resp. $m$) appelé *majorant* (resp. *minorant*) tel que $x_k \leqslant M$ (resp. $x_k \geqslant m$) pour tous les $k = 0, 1, 2, \ldots$

Les termes de la suite (1) étant entiers, on dira que cette suite *se stabilise* à un nombre $\xi$ si l'on peut exhiber un $k_0$ tel que $x_k = \xi$ pour tous les $k > k_0$ et on notera $x_k \underset{\rightarrow}{\rightarrow} \xi$.

Lemme 1. *Si une suite d'entiers est croissante et majorée par un nombre $M$, elle se stabilise à un entier $\xi \leqslant M$.*

Démonstration. La suite $\{x_n\}$ est majorée par le nombre $M$ et est croissante, donc elle est minorée. Bien qu'elle comprenne un nombre infini de termes, cette suite ne parcourt qu'un nombre *fini* d'entre eux, car ils sont $\leqslant M$. Soit $\xi$ le plus grand de ces termes. On a $\xi \leqslant M$ et il existe un indice $s$ tel que $x_s = \xi$. Or, la suite $\{x_n\}$ est croissante, donc $x_k = x_s$ pour tous les $k \geqslant s$, autrement dit la suite se stabilise au nombre $\xi$ ($x_n \underset{\rightarrow}{\rightarrow} \xi \leqslant M$).

Considérons maintenant une suite de formes décimales positives (simultanément limitées ou illimitées):

$$\left\{\begin{array}{l} a_1 = \alpha_{10}, \ \alpha_{11}\alpha_{12}\alpha_{13} \ldots, \\ a_2 = \alpha_{20}, \ \alpha_{21}\alpha_{22}\alpha_{23} \ldots, \\ a_3 = \alpha_{30}, \ \alpha_{31}\alpha_{32}\alpha_{33}, \ldots \\ \ldots\ldots\ldots\ldots\ldots \end{array}\right. \tag{2}$$

Les seconds membres de (2) forment un tableau (*matrice infinie*).

On dira que la *suite* (2) *se stabilise au nombre* $a = \gamma_0, \gamma_1\gamma_2 \ldots$, et on écrira

$$a_n \underset{\rightarrow}{\rightarrow} a, \tag{3}$$

*si la k-ième colonne du tableau* (2) *se stabilise à* $\gamma_k$ *quel que soit* $k =$ $= 0, 1, 2, \ldots$, c'est-à-dire que $\alpha_{sk} \underset{\rightarrow}{\rightarrow} \gamma_k$ pour tout $k$ fixe.

Lemme 2. *Si une suite croissante* (2) *de formes décimales (toutes limitées ou toutes illimitées) est majorée par un nombre $M$, elle se stabilise à fortiori à un nombre $a$ tel que*

$$a_n \leqslant a \leqslant M \quad (n = 1, 2, \ldots). \tag{4}$$

En effet, dans les hypothèses du lemme, les entiers de la colonne zéro de la matrice (2)

$$\begin{matrix} \alpha_{10} \\ \alpha_{20} \\ \alpha_{30} \\ \cdots \end{matrix}$$

croissent aussi et sont majorés par le nombre $M$ (cf. § 5), donc en vertu du lemme 1, ils se stabilisent à un entier positif $\gamma_0 \leqslant M$. Supposons que cette stabilisation a lieu à partir d'un indice $n_0$, c'est-à-dire que $a_n = \gamma_0, \alpha_{n1}\alpha_{n2} \ldots \leqslant M, \ n \geqslant n_0$.

Prouvons maintenant que la première colonne de (2)

$$\begin{matrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \\ \cdots \end{matrix}$$

se stabilise aussi à un chiffre $\gamma_1$ et que

$$\gamma_0, \gamma_1 \leqslant M.$$

En effet, les représentations décimales des nombres $a_n$ étant pour $n \geqslant n_0$ de la forme

$$\gamma_0, \alpha_{n1}\alpha_{n2}\alpha_{n3} \ldots \leqslant M$$

et la suite $\{a_n\}$ étant croissante, les chiffres $\alpha_{n1}$ ($\leqslant 9$) de la première colonne croissent pour $n \geqslant n_0$ et par suite se stabilisent, en vertu du lemme 1, à un chiffre $\gamma_1$. Supposons que la deuxième stabilisation a lieu à partir de l'indice $n_1 > n_0$, autrement dit pour $n \geqslant n_1$,

$$a_n = \gamma_0, \gamma_1\alpha_{n2}\alpha_{n3} \ldots \leqslant M.$$

Ceci étant, il est manifeste que

$$\gamma_0, \gamma_1 \leqslant a_n \leqslant M \quad (n \geqslant n_1).$$

Raisonnons maintenant par récurrence. Supposons que les colonnes de la matrice (2) d'indices inférieurs à $k$ se stabilisent respectivement à $\gamma_0, \gamma_1, \ldots, \gamma_k$ et que

$$\gamma_0, \gamma_1 \ldots \gamma_k \leqslant M \ (\gamma_1, \ldots, \gamma_k \text{ sont des chiffres}). \tag{5}$$

Prouvons que la $(k+1)$-ième colonne de (2) se stabilise aussi à un chiffre $\gamma_{k+1}$ et que

$$\gamma_0, \gamma_1 \ldots \gamma_k\gamma_{k+1} \leqslant M. \tag{6}$$

En effet, puisque les représentations décimales des nombres $a_n$ sont, pour $n \geqslant n_k$, de la forme

$$a_n = \gamma_0, \gamma_1 \ldots \gamma_k\alpha_{n,k+1}\alpha_{n,k+2} \ldots \leqslant M$$

et de plus $a_n$ croît, les chiffres $\alpha_{n,k+1}$ croissent pour $n \geqslant n_k$ et par

suite se stabilisent pour $n \geqslant n_{k+1}$, où $n_{k+1}$ est assez grand, à un chiffre $\gamma_{k+1}$. De plus il est évident que

$$\gamma_0,\ \gamma_1 \ldots \gamma_{k+1} \leqslant a_n \leqslant M \quad (n \geqslant n_{k+1}),$$

c'est-à-dire qu'on retrouve l'inégalité (6). Nous avons donc prouvé par récurrence que (5) est vraie pour tout $k$ et que $a_n \rightrightarrows a = \gamma_0,\ \gamma_1\gamma_2 \ldots$

Prouvons la première inégalité (4). Comparons les nombres

$$a_n = \alpha_{n0},\ \alpha_{n1}\alpha_{n2}\alpha_{n3} \ldots$$

et

$$a = \gamma_0,\ \gamma_1\gamma_2\gamma_3 \ldots$$

Si les composantes respectives des deux représentations sont égales ($\alpha_{ns} = \gamma_s$, $s = 0, 1, 2, \ldots$), alors $a_n = a$. Sinon, pour un certain $s$ on a

$$\left\{\begin{array}{l} \alpha_{nj} = \gamma_j \quad (j = 0, 1, \ldots, s-1), \\ \alpha_{ns} < \gamma_s. \end{array}\right. \tag{7}$$

Ceci étant, si $s = 0$, les égalités n'ont pas de sens dans (7). Pour l'indice $n$ considéré, les nombres $\alpha_{nj} = \gamma_j$ ($j = 0, 1, \ldots, s-1$) sont déjà stabilisés, donc

$$a_n \leqslant \alpha_{n0},\ \alpha_{n1} \ldots \alpha_{n,s-1}(\alpha_{ns}+1) = \gamma_0,\ \gamma_1 \ldots \gamma_{s-1}(\alpha_{ns}+1) \leqslant \leqslant \gamma_0,\ \gamma_1 \ldots \gamma_{s-1}\gamma_s \leqslant a,$$

ce qui prouve la première inégalité (4).

Reste à démontrer la deuxième inégalité (4). Si $a = \gamma_0,\ \gamma_1 \ldots \gamma_N$ est une forme décimale limitée, alors elle résulte de (5) pour $k = N$. Supposons maintenant que

$$a = \gamma_0,\ \gamma_1\gamma_2 \ldots \tag{8}$$

est une forme décimale illimitée. Représentons le nombre $M$ par une forme décimale illimitée $M = m_0,\ m_1m_2 \ldots$ Si l'on admet que l'inégalité à démontrer est fausse, il existe alors un $k$ tel que

$$\left\{\begin{array}{l} \gamma_j = m_j \quad (j = 0, 1, \ldots, k-1), \\ \gamma_k > m_k. \end{array}\right. \tag{9}$$

Si $k = 0$, les égalités n'ont pas de sens dans (9). La forme (8) étant illimitée, il existe un $s$ tel que $\gamma_{k+s} > 0$. Donc

$$\gamma_0,\ \gamma_1 \ldots \gamma_{k-1}\gamma_k \ldots \gamma_{k+s} > \gamma_0,\ \gamma_1 \ldots \gamma_{k-1}\gamma_k = m_0,\ m_1 \ldots m_{k-1}\gamma_k \geqslant \geqslant m_0,\ m_1 \ldots m_{k-1}(m_k+1) = m_0, m_1 \ldots m_{k-1}m_k99 \ldots \geqslant m_0, m_1m_2 \ldots = M,$$

ce qui contredit l'inégalité (5).

Nous sommes désormais en mesure de définir les opérations arithmétiques sur les nombres réels.

Considérons un nombre quelconque $a = \alpha_0,\ \alpha_1\alpha_2 \ldots$ et introduisons sa *n-troncature* $a^{(n)} = \alpha_0,\ \alpha_1\alpha_2 \ldots \alpha_n$. On admet que le lecteur connaît les opérations sur les formes décimales limitées.

Soient deux nombres strictement positifs $a$ et $b$ dont les formes décimales illimitées s'écrivent

$$a = \alpha_0,\ \alpha_1\alpha_2 \ldots, \qquad b = \beta_0,\ \beta_1\beta_2 \ldots$$

Introduisons la suite de nombres

$$a^{(n)} + b^{(n)} = \alpha_0,\ \alpha_1 \ldots \alpha_n + \beta_0,\ \beta_1 \ldots \beta_n = \lambda_0^{(n)},\ \lambda_1^{(n)} \ldots \lambda_n^{(n)}$$
$$(n = 1, 2, \ldots).$$

Il est évident que cette suite est croissante et qu'elle est majorée:

$$a^{(n)}+b^{(n)}\leqslant(\alpha_0+1)+(\beta_0+1)\quad (n=1,\ 2,\ \ldots).$$

Donc, en vertu du lemme 2, les représentations décimales de ses éléments se stabilisent à une forme décimale $\gamma_0,\ \gamma_1\gamma_2\ldots$ correspondant à un nombre réel. Ce nombre est par définition la *somme des nombres* $a$ et $b$:

$$a+b=\gamma_0,\ \gamma_1\gamma_2\ldots$$

Ainsi nous définissons la somme $a+b$ comme le nombre auquel se stabilise $\{a^{(n)}+b^{(n)}\}$:

$$a^{(n)}+b^{(n)}\rightrightarrows a+b. \tag{10}$$

Pour définir le produit des nombres positifs $a$ et $b$, introduisons la troncature

$$(a^{(n)}b^{(n)})^{(n)}=\mu_0^{(n)},\ \mu_1^{(n)}\ldots\mu_n^{(n)}. \tag{11}$$

La suite de ces troncatures est visiblement croissante (lorsque $n$ croît) et est majorée:

$$(a^{(n)}b^{(n)})^{(n)}\leqslant(\alpha_0+1)(\beta_0+1)\quad (n=1,\ 2,\ \ldots).$$

Donc, le lemme 2 nous dit que cette suite se stabilise à un nombre que l'on appellera *produit* $ab$:

$$(a^{(n)}b^{(n)})^{(n)}\rightrightarrows ab.$$

Signalons les inégalités

$$a^{(n)}=\alpha_0,\alpha_1\ldots\alpha_n\leqslant\alpha_0,\alpha_1\ldots\alpha_n\ldots\leqslant\alpha_0,\alpha_1\ldots\alpha_n99\ldots=$$
$$=\alpha_0,\ \alpha_1\ldots\alpha_{n-1}(\alpha_n+1)=\alpha_0,\ \alpha_1\ldots\alpha_n+10^{-n},$$

i.e.

$$a^{(n)}\leqslant a\leqslant a^{(n)}+10^{-n}.$$

La quantité $a^{(n)}$ tend vers $a$ (lorsque $n$ croît) en croissant. Quant à la quantité $a^{(n)}+10^{-n}$, elle tend vers $a$ en décroissant:

$$a^{(n)}+10^{-n}=\alpha_0,\ \alpha_1\ldots\alpha_n99\ldots\geqslant\alpha_0,\ \alpha_1\ldots\alpha_n\alpha_{n+1}99\ldots=$$
$$=a^{(n+1)}+10^{-(n+1)}.$$

Cette circonstance sera utilisée pour définir la différence et le quotient de nombres strictement positifs.

Si $a>b>0$, la *différence* $a-b$ se définit comme une forme décimale à laquelle se stabilise une suite de formes décimales limitées:

$$a^{(n)}-(b^{(n)}+10^{-n})\rightrightarrows(a-b); \tag{12}$$

si $a,\ b>0$, alors le *quotient* $a\ /\ b$ se définit comme une forme décimale à laquelle se stabilise la suite des formes décimales limitées:

$$\left(\frac{a^{(n)}}{b^{(n)}+10^{-n}}\right)^{(n)}\rightrightarrows\frac{a}{b}. \tag{13}$$

A noter que lorsque $n$ croît, $a^{(n)}$ croît et $b^{(n)} + 10^{-n}$ décroît, donc les premiers membres de (12) et (13) croissent et de plus ils sont majorés:

$$a^{(n)} - (b^{(n)} + 10^{-n}) \leqslant \alpha_0 + 1,$$

$$\left(\frac{a^{(n)}}{b^{(n)} + 10^{-n}}\right)^{(n)} \leqslant \frac{\alpha_0 + 1}{\beta_0, \beta_1 \ldots \beta_s},$$

où $s$ est tel que $\beta_s > 0$. Donc, en vertu du lemme 2, ils se stabilisent.

Posons encore

$$0 + a = a \pm 0 = a, \quad a \cdot 0 = 0 = a - a = \frac{0}{b} \quad (a \geqslant 0, \ b > 0). \qquad (14)$$

Nous avons défini la somme, la différence, le produit et le quotient de nombres positifs $a$, $b$ en admettant dans le cas de la différence que $a \geqslant b$ et dans le cas du quotient que $b > 0$. Ces définitions se généralisent par les procédés habituels aux nombres $a$ et $b$ de signe quelconque. Par exemple, si $a$, $b \leqslant 0$, on convient que $a + b = b + a = -(|a| + |b|)$. Si $a$ et $b$ sont de signes contraires et $|a| \geqslant |b|$, on admet que $a + b = b + a = \pm(|a| - |b|)$, le signe retenu étant celui de $a$. En particulier

$$a + (-a) = 0$$

quel que soit $a$.

Les autres opérations sont justiciables de règles identiques que nous ne citerons pas, car elles sont bien connues du cours d'algèbre.

Au § 7, on énumère les propriétés des nombres réels découlant des définitions ci-dessus. Ces propriétés ne seront démontrées que dans certains cas [1]). Elles ont été réparties en cinq groupes (I à V). Les trois premiers groupes comprennent les propriétés élémentaires dont on se sert pour les calculs arithmétiques et la résolution des inégalités. Le groupe IV est composé d'une propriété (la propriété d'Archimède). Le groupe V enfin est constitué d'une propriété formulée en termes de limites qui sera prouvée au § 5, chap. 2.

## § 7. Propriétés fondamentales des nombres réels

I. Relation d'ordre.

$I_1$. *Axiome de trichotomie. Pour tout couple de nombres réels $a$ et $b$, on a*

$$a = b \quad ou \quad a > b \quad ou \quad a < b,$$

*ces cas s'excluant.*

$I_2$. *De $a < b$ et $b < c$ il s'ensuit que $a < c$ (la relation est transitive).*

$I_3$. *Si $a < b$, il existe un nombre $c$ tel que $a < c < b$.*

[1]) La démonstration complète figure dans l'ouvrage de Nikolski « Analyse mathématique », Tome I, chap. 2.

II. Propriétés de l'addition et de la soustraction.

$\mathrm{II}_1$. $a + b = b + a$ (*commutativité*).

$\mathrm{II}_2$. $(a + b) + c = a + (b + c)$ (*associativité*).

$\mathrm{II}_3$. $a + 0 = a$.

$\mathrm{II}_4$. $a + (-a) = 0$.

$\mathrm{II}_5$. *De* $a < b$ *il résulte que* $a + c < b + c$ *pour tout* $c$.

Le nombre $a + (-b)$ s'appelle naturellement *différence* entre $a$ et $b$, c'est-à-dire que $a - b = a + (-b)$, car si on ajoute cette différence à $b$, on obtient le nombre $a$:

$$[a + (-b)] + b = a + [(-b) + b] = a + 0 = a.$$

On voit sans peine à partir des propriétés précédentes que la différence est unique. On montre que la différence telle qu'elle vient d'être définie coïncide avec celle donnée par la formule (12), § 6.

III. Propriétés de la multiplication et de la division.

$\mathrm{III}_1$ $ab = ba$ (*commutativité*).

$\mathrm{III}_2$. $(ab)\,c = a\,(bc)$ (*associativité*).

$\mathrm{III}_3$. $a \cdot 1 = a$.

$\mathrm{III}_4$. $a \cdot \frac{1}{a} = 1$ $(a \neq 0)$.

$\mathrm{III}_5$. $(a + b)\,c = ac + bc$ (*distributivité*).

$\mathrm{III}_6$. *De* $a < b$ *et* $c > 0$ *il s'ensuit que* $ac < bc$.

Il est naturel d'appeler le nombre $a \cdot \frac{1}{b}$ $(b \neq 0)$ *quotient* de $a$ par $b$ $\left(\frac{a}{b} = a \cdot \frac{1}{b}\right)$, car en le multipliant par $b$ on obtient $a$:

$$\left(a \cdot \frac{1}{b}\right) b = a \left(\frac{1}{b} \cdot b\right) = a \left(b \cdot \frac{1}{b}\right) = a \cdot 1 = a.$$

Il résulte des propriétés précédentes que le quotient de $a$ par $b$ est unique. On démontre que tel qu'il vient d'être défini le quotient coïncide avec celui donné par la formule (13), § 6.

IV. Propriété d'Archimède.

*Pour tout nombre* $c > 0$, *il existe un entier naturel* $n > c$. En effet, si $c = \alpha_0, \alpha_1\alpha_2 \ldots$, on peut prendre $n = \alpha_0 + 2$.

De la propriété d'Archimède et de certaines autres il s'ensuit que pour tout nombre strictement positif $\varepsilon$ on peut exhiber un entier naturel $n$ tel que $\frac{1}{n} < \varepsilon$.

En effet, en vertu de IV on peut exhiber pour le nombre $1/\varepsilon$ un entier naturel $n$ tel que $1/\varepsilon < n$, ce qui d'après $\mathrm{III}_6$ entraîne l'inégalité désirée.

On remarquera que pour un nombre donné $c \geqslant 0$ il existe dans la suite des entiers positifs $0, 1, 2, \ldots$ un seul nombre $m$ tel que $m \leqslant c < m + 1$.

P r o p r i é t é V. *Si une suite de nombres réels* $a_1, a_2, a_3 \dots$ *est croissante et majorée par un nombre* $M$ ($a_n \leqslant M$), *il existe alors un nombre* $a \leqslant M$ *qui est la limite de cette suite*:

$$\lim_{n\to\infty} a_n = a \leqslant M.$$

Cela signifie que pour tout nombre $\varepsilon > 0$ aussi petit que l'on veut on peut exhiber un entier naturel $n_0$ tel que

$$|a - a_n| = a - a_n < \varepsilon$$

pour tous les $n > n_0$.

Voir démonstration du théorème. § 5, chap. 2. Nous verrons que la propriété V est une conséquence immédiate du lemme 2, § 6, qui affirme, en particulier, qu'une suite de formes décimales illimitées croissante majorée par un nombre $M$ se stabilise à une forme décimale représentant un nombre $a \leqslant M$ ($a_n \rightrightarrows a$).

En effet, dire que $a_n$ se stabilise à $a$ revient à dire que $a$ est la limite de $a_n$.

## § 8. Approche axiomatique de la notion de nombre réel

Nous avons appelé nombres réels les formes décimales illimitées et avons introduit pour eux les notions de 0, 1, $>$, $=$, les opérations arithmétiques. Nous avons formulé les propriétés I à V.

Signalons qu'à partir des propriétés I à V on peut déduire toutes les autres propriétés des nombres.

L'approche axiomatique consiste à définir les nombres réels comme des objets (êtres) $a$, $b$, $c$, . . . satisfaisant aux propriétés I à V. Dans cette approche les propriétés I à V s'appellent *axiomes*.

Les axiomes doivent être légèrement modifiés. L'axiome II s'énonce dorénavant: *à chaque couple de nombres est associé un nombre* $a + b$ *appelé somme et vérifiant les axiomes* $\mathrm{II}_1$ *à* $\mathrm{II}_5$. L'axiome $\mathrm{II}_3$ devient: *il existe un nombre* 0 (*zéro*) *tel que* $a + 0 = a$ *pour tous les* $a$. L'axiome $\mathrm{II}_4$ prend la forme: *pour tout nombre* $a$ *il existe un nombre* $-a$ *tel que* $a + (-a) = 0$. Enfin, l'axiome $\mathrm{III}_3$ se formule: *il existe un nombre* 1 (*unité*) *distinct de* 0 *et tel que* $a \cdot 1 = a$ *pour tous les* $a$.

Désignons par **R** l'ensemble de tous les nombres réels, c'est-à-dire de tous les êtres vérifiant les axiomes I à V. Donc, **R** contient les éléments 0 et 1. Les axiomes nous permettent de prouver que $0 < 1$ et que les nombres $2 = 1 + 1$, $3 = 2 + 1$, . . ., et $-1$, $-2$, $-3$, . . . ont un sens. On obtient en définitive l'ensemble des entiers relatifs (distincts entre eux!)

$$\dots, -2, -1, 0, 1, 2, \dots$$

En vertu de ces axiomes la division des réels est possible sauf par zéro. Donc, l'ensemble **R** comprend les nombres rationnels

$\pm m/n = \pm mp/np$ ($n > 0$, $m \geqslant 0$, $p \neq 0$). Il contient également les formes décimales limitées. Avec ces dernières on peut construire des suites majorées croissantes. L'axiome V nous dit que ces suites admettent des limites dans **R**. Certaines de ces limites ne sont pas des formes décimales limitées, mais des nombres distincts d'elles qu'il est commode de représenter par des formes décimales illimitées. En partant des axiomes on a abouti par des raisonnements logiques aux formes décimales illimitées. Nous avons esquissé la marche à suivre mais celle-ci ne prétend nullement à une démonstration.

De ce qui précède il s'ensuit que d'un point de vue formel peu importe que l'on définisse les nombres réels à partir des formes décimales illimitées ou axiomatiquement.

Du point de vue philosophique, la seconde approche est certes mieux appropriée dans la mesure où les nombres sont des êtres abstraits exprimant les rapports quantitatifs du monde réel, alors que les formes décimales ne sont que des symboles formels les représentant.

## § 9. Inégalités pour les valeurs absolues

L'inégalité

$$|a| < \varepsilon \tag{1}$$

équivaut aux deux inégalités

$$-\varepsilon < a < \varepsilon. \tag{1'}$$

De là, l'inégalité

$$|a - b| < \varepsilon \tag{2}$$

équivaut aux inégalités

$$b - \varepsilon < a < b + \varepsilon. \tag{2'}$$

De façon analogue, l'inégalité

$$|a - b| \leqslant \varepsilon \tag{3}$$

équivaut aux inégalités

$$b - \varepsilon \leqslant a \leqslant b + \varepsilon.$$

On a de même

$$|a + b| \leqslant |a| + |b|, \tag{4}$$

$$|a - b| \geqslant ||a| - |b||. \tag{5}$$

On établit l'inégalité (4) en distinguant quatre cas: 1) $a, b \geqslant 0$, 2) $a, b \leqslant 0$, 3) $a \leqslant 0 \leqslant b$, 4) $b \leqslant 0 \leqslant a$. Dans le cas 2) par exemple

$$a + b \leqslant b \leqslant 0, \quad |a + b| = -(a + b) = -a - b = \\ = |a| + |b|$$

dans le cas 3) si l'on admet que $|b| \geqslant |a|$

$$|a + b| = b + a \leqslant |a| + |b|.$$

Le cas 3) pour $|b| \leqslant |a|$, de même que le cas 1) sont laissés au soin du lecteur. Le cas 4) se ramène au cas 3). D'autre part, en vertu de (4) on a

$$|a| \leqslant |b| + |a - b|, \quad |b| \leqslant |a| + |a - b|,$$

c'est-à-dire que

$$|a| - |a - b| \leqslant |b| \leqslant |a| + |a - b|,$$

d'où (5).

## § 10. Intervalles, ensemble borné

Soient $a$ et $b$ des nombres tels que $a < b$.

L'ensemble des nombres $x$ tels que $a \leqslant x \leqslant b$ est un intervalle appelé *intervalle fermé* d'origine $a$ et d'extrémité $b$ et noté $[a, b]$.

L'ensemble des nombres $x$ tels que $a < x < b$ est un intervalle appelé *intervalle ouvert* d'origine $a$ et d'extrémité $b$ et noté $]a, b[$.

L'ensemble des nombres $x$ tels que $a \leqslant x < b$ (resp. $a < x \leqslant b$) est un *intervalle semi-ouvert à droite* (resp. *à gauche*) d'origine $a$ et d'extrémité $b$ et noté $[a, b[$ (resp. $]a, b]$).

Les ensembles des nombres $x$ tels que: 1) $x \leqslant a$; 2) $x < a$; 3) $x \geqslant a$; 4) $x > a$ sont des intervalles appelés respectivement 1) *intervalle fermé illimité à gauche* (ou section commençante fermée d'extrémité $a$); 2) *intervalle ouvert illimité à gauche* (ou section commençante ouverte d'extrémité $a$); 3) *intervalle fermé illimité à droite* (ou section finissante fermée d'origine $a$); 4) *intervalle ouvert illimité à droite* (ou section finissante ouverte d'origine $a$), et notés 1) $]-\infty, a]$; 2) $]-\infty, a[$, $[a, \infty[$, $]a, \infty[$. L'ensemble **R** se note aussi $]-\infty, \infty[$.

Les symboles $-\infty$ et $+\infty$ représentent les *nombres infinis*.

Signalons que les extrémités de l'intervalle fermé $[a, b]$ sont des nombres finis, alors que celles de l'intervalle ouvert $]a, b[$ peuvent être aussi bien finies qu'infinies. De même l'extrémité $a$ (resp. $b$) de l'intervalle $[a, b[$ (resp. $]a, b]$) est toujours un nombre fini, tandis que l'extrémité $b$ (resp. $a$) peut aussi bien être finie qu'infinie.

Si $a$ et $b$ sont finis et $a < b$, alors le nombre $b - a$ s'appelle *longueur* de l'intervalle fermé $[a, b]$, ou ouvert $]a, b[$, ou semi-ouvert à gauche $]a, b]$, ou semi-ouvert à droite $[a, b[$.

On appellera *voisinage d'un point* $c$ tout intervalle ouvert $]a, b[$ contenant le point $c$ $(a < c < b)$. En particulier, l'intervalle $]c - \varepsilon, c + \varepsilon[$ $(\varepsilon > 0)$ est dit *ε-voisinage du point* $c$.

Soit $X = \{x\}$ un ensemble de nombres réels. On dit que l'ensemble $X$ est *majoré* s'il existe un nombre $M$ tel que pour tout $x \in X$ l'on ait $x \leqslant M$; *minoré* s'il existe un nombre $m$ tel que pour tout $x \in X$ l'on ait $x \geqslant m$; *borné* s'il est à la fois majoré et minoré.

De toute évidence, il revient au même de dire qu'un ensemble est borné si $\exists\, M > 0, \forall\, x \in X: |x| \leqslant M$, car l'inégalité $|x| \leqslant M$ équivaut à la double inégalité $-M \leqslant x \leqslant M$.

Si l'ensemble $X$ n'est pas borné, on dit qu'il est *non borné*. On peut alors le définir comme suit : un ensemble $X$ de nombres réels est non borné si $\forall M > 0, \exists x_0 \in X : |x_0| > M$. On aurait pu accéder à cet énoncé par négation de la formule logique.

EXEMPLES. L'intervalle $[a, b]$ est un ensemble borné. L'intervalle $]a, b[$ est un ensemble borné si $a$ et $b$ sont finis et non borné si $a = -\infty$ ou $b = \infty$.

## § 11. Ensemble dénombrable. Dénombrabilité de l'ensemble des rationnels. Non-dénombrabilité de l'ensemble des réels

Nous avons défini plus haut la notion d'égalité d'ensembles. La notion d'équivalence est commode pour la caractérisation du degré de saturation des ensembles infinis. On dit qu'un ensemble $X$ est *infini* si $\forall n \in \mathbf{N}$, le nombre des éléments de $X$ est supérieur à $n$. Deux ensembles $A$ et $B$ sont dits *équivalents* ($A \sim B$) si l'on peut mettre leurs éléments en correspondance biunivoque ($\longleftrightarrow$), c'est-à-dire s'il existe une loi qui à tout $a \in A$ associe un élément $b \in B$ et un seul et inversement à tout $b \in B$ est associé un élément $a \in A$ et un seul.

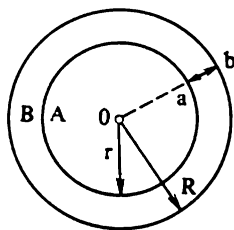


Fig. 6

Si par exemple $A$ est l'ensemble des points d'un cercle de rayon $r$, $B$ celui des points d'un cercle concentrique de rayon $R > r$, alors il est manifeste que $A \sim B$ (fig. 6).

Il est évident que si $A = B$, alors $A \sim B$.

On dit qu'un ensemble est *dénombrable* s'il est équivalent à l'ensemble des entiers naturels $\mathbf{N}$. L'ensemble $\mathbf{N}$ est dénombrable ($n \longleftrightarrow n$). L'ensemble $\mathbf{N}_p = \{2n\}$ des entiers naturels pairs est équivalent à l'ensemble $\mathbf{N}$ tout entier ($n \longleftrightarrow 2n$). Signalons que $\mathbf{N}_p \neq \mathbf{N}$, $\mathbf{N}_p \subset \mathbf{N}$. On voit donc qu'un ensemble est équivalent à l'une de ses parties. Cette propriété n'est caractéristique que des ensembles infinis (et peut être adoptée pour définition d'un ensemble infini).

De la définition de la dénombrabilité d'un ensemble il résulte qu'on peut numéroter les éléments d'un ensemble dénombrable avec les entiers naturels, donc on représentera souvent un tel ensemble par la suite de ses éléments :

$$X = \{x_1, x_2, \ldots, x_n, \ldots\}.$$

*La somme dénombrable*

$$E = \bigcup_{k=1}^{\infty} E^k = E^1 + E^2 + \ldots$$

*d'ensembles dénombrables $E^k$ est un ensemble dénombrable.* En effet, écrivons les éléments $x_j^k \in E^k$ ($j = 1, 2, \ldots$) sous forme de tableau:

$$E^1 = \{x_1^1, x_2^1, x_3^1, \ldots\},$$
$$E^2 = \{x_1^2, x_2^2, x_3^2, \ldots\},$$
$$E^3 = \{x_1^3, x_2^3, x_3^3, \ldots\},$$
$$\cdots\cdots\cdots$$

Rangeons-les dans l'ordre suivant:

$$x_1^1,\ x_2^1,\ x_1^2,\ x_3^1,\ x_2^2,\ x_1^3,\ x_4^1,\ \ldots$$

Si des ensembles $E^k$ et $E^l$ présentent des éléments communs, on ne gardera qu'un seul d'entre eux. On obtient en définitive une suite infinie d'éléments qui épuise visiblement l'ensemble $E$. Ceci prouve que $E$ est un ensemble dénombrable.

On démontre de façon analogue qu'*une somme finie $E = E^1 + \ldots + E^N$ d'ensembles finis et d'au moins un ensemble dénombrable est un ensemble dénombrable.*

THÉORÈME 1. *L'ensemble* $\mathbf{Q}$ *des nombres rationnels est dénombrable.*

DÉMONSTRATION. Considérons d'abord l'ensemble $\mathbf{Q}_+ = \{p/q\}$ des nombres rationnels strictement positifs. Appelons le nombre $p + q$ *hauteur* du nombre rationnel $p/q$. Soit $A_n$ l'ensemble de tous les nombres rationnels de hauteur $n$. Les ensembles $A_n$ sont composés d'un nombre fini d'éléments. Par exemple

$$A_1 = \varnothing,\quad A_2 = \left\{\frac{1}{1}\right\},\quad A_3 = \left\{\frac{1}{2}, \frac{2}{1}\right\},\quad A_4 = \left\{\frac{1}{3}, \frac{2}{2}, \frac{3}{1}\right\}, \ldots$$

Il est aisé de voir que $\mathbf{Q}_+ = \bigcup\limits_{n=1}^{\infty} A_n$.

Renumérotons les nombres entre accolades à partir de la gauche en ne gardant que ceux qui apparaissent une seule fois. On obtient en définitive la suite

$$r_1 = 1,\quad r_2 = \frac{1}{2},\quad r_3 = 2,\quad r_4 = \frac{1}{3},\quad r_5 = 3, \ldots$$

Donc, l'ensemble $\mathbf{Q}_+$ est dénombrable. Il est évident par ailleurs que $\mathbf{Q}_- = \{-p/q\}$ est dénombrable. Donc, l'ensemble des nombres rationnels $\mathbf{Q} = \mathbf{Q}_+ \cup \mathbf{Q}_- \cup \{0\}$ est aussi dénombrable.

THÉORÈME 2. *L'ensemble* $\mathbf{R}$ *des nombres réels n'est pas dénombrable.*

DÉMONSTRATION. Il suffit de prouver que l'ensemble des nombres réels de l'intervalle $]0, 1[$ est non dénombrable. Supposons par absurde que l'intervalle

] 0, 1 [ est un ensemble dénombrable, autrement dit qu'on peut numéroter tous ses points:

$$x_1 = 0,\ a_1^{(1)} a_2^{(1)} \ldots,$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$
$$x_n = 0,\ a_1^{(n)} a_2^{(n)} \ldots,$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

Considérons le nombre réel $x = 0,\ a_1 a_2 \ldots$, où $0 < a_n < 9$ et $a_n \neq a_n^{(n)}$. Il est clair que $x \in\ ]0, 1[$ et qu'il n'est confondu avec aucun $x_n$. Donc aucune suite $x_1, \ldots, x_n, \ldots$ ne peut épuiser les nombres réels compris entre 0 et 1.

CHAPITRE 2

# LIMITE D'UNE SUITE

## § 1. Notion de limite d'une suite

Supposons qu'une loi associe à tout entier naturel $n = 1, 2, 3, \ldots$ un nombre réel ou complexe $x_n$.

On vient ainsi de définir une *suite de nombres* $x_1, x_2, x_3, \ldots$ ou tout simplement la suite

$$\{x_n\} = \{x_1, x_2, x_3, \ldots\}.$$

On dit encore que la variable $x_n$ parcourt les valeurs de la suite $\{x_n\}$.

Les nombres $x_n$ s'appellent *éléments* ou *termes* de la suite $\{x_n\}$. Si $n \neq m$, les éléments $x_n$ et $x_m$ sont considérés comme distincts même s'ils sont égaux.

Dans ce chapitre on examinera les suites de nombres *réels* sans le spécifier expressément.

Voici quelques exemples de suites:

EXEMPLE 1. $\left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\} = \left\{\frac{1}{n}\right\}$.

EXEMPLE 2. $\left\{\frac{1}{2}, 2, \frac{1}{2}, 2, \ldots\right\} = \{2^{(-1)^n}\}$.

EXEMPLE 3. $\left\{1, 2, \frac{1}{3}, 4, \frac{1}{5}, \ldots\right\} = \{n^{(-1)^n}\}$.

EXEMPLE 4. $\left\{0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\right\} = \left\{\frac{n-1}{n}\right\}$.

EXEMPLE 5. $\{2, 5, 10, \ldots\} = \{n^2 + 1\}$.

EXEMPLE 6. $\{-1, 2, -3, 4, \ldots\} = \{(-1)^n n\}$.

Dans l'exemple 2, la variable $x_n$ prend la même valeur pour $n$ pair:

$$2 = x_2 = x_4 = x_6 = \ldots$$

On considère cependant que les éléments $x_2, x_4, \ldots$ sont différents.

Si tous les termes d'une suite $\{x_n\}$ sont égaux à un même nombre $a$, on dit que cette suite est *constante*.

Il est immédiat de voir que les suites des exemples 1, 2 et 4 sont *bornées* (voir § 6, chap. 1). On dit encore que les variables parcourant ces suites sont bornées. Les suites des exemples 3, 5 et 6 ne sont pas bornées. La suite de l'exemple 3 est minorée par 0, celle de l'exemple 5, par le nombre 2. La suite de l'exemple 6 n'est ni minorée ni majorée.

Introduisons la notion de limite d'une suite.

DÉFINITION. *On dit qu'un nombre $a$ est la limite d'une suite $\{x_n\}$ si pour tout $\varepsilon > 0$ il existe un nombre $n_0 = n_0(\varepsilon)$ dépendant de $\varepsilon$ tel que*

$$|x_n - a| < \varepsilon \tag{1}$$

*pour tous les $n > n_0$.*

On écrit

$$\lim_{n\to\infty} x_n = \lim x_n = a \quad \text{ou} \quad x_n \to a$$

et on dit que la variable $x_n$ ou la suite $\{x_n\}$ a le nombre $a$ pour limite ou tend vers $a$ ou encore *converge vers $a$.*

Si $x_n = a$, $\forall n \in \mathrm{N}$, il est évident que $\lim\limits_{n\to\infty} x_n = \lim a = a$.

REMARQUE 1. Si $\lim x_n = a$, alors $\lim x_{n+1} = a$, et inversement. Ceci résulte du fait que si

$$|x_n - a| < \varepsilon, \quad \forall n > n_0,$$

alors

$$|x_{n+1} - a| < \varepsilon, \quad \forall n > n_0 - 1,$$

et inversement.

La variable de l'exemple 1 tend vers 0:

$$\lim_{n\to\infty} \frac{1}{n} = 0. \tag{2}$$

En effet, considérons un nombre arbitraire $\varepsilon > 0$ et résolvons l'inégalité

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon \quad \text{ou} \quad \frac{1}{\varepsilon} < n.$$

Donc, pour tout $\varepsilon > 0$ on a trouvé un nombre $n_0 = n_0(\varepsilon) = 1/\varepsilon$ tel que l'inégalité

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon$$

soit réalisée pour tous les $n > n_0$. Ceci prouve (2).

EXEMPLE 7. La suite de l'exemple 4 tend vers 1:

$$\lim_{n\to\infty} \frac{n-1}{n} = 1. \tag{3}$$

En effet, formons l'inégalité

$$\left|1-\frac{n-1}{n}\right|=\frac{1}{n}<\varepsilon.$$

On a vu qu'elle était réalisée pour tout $\varepsilon>0$ si $n>n_0=1/\varepsilon$. Ceci prouve (3).

EXEMPLE 8. Si $|q|<1$, alors

$$\lim_{n\to\infty} q^n=0. \tag{4}$$

En effet, supposons que $q\neq 0$. L'inégalité

$$|q^n-0|=|q^n|<\varepsilon$$

est vérifiée si

$$n\log|q|<\log\varepsilon,$$

c'est-à-dire si

$$n>\frac{\log\varepsilon}{\log|q|}=n_0(\varepsilon).$$

Nous avons prouvé (4) pour $0<|q|<1$. Si $q=0$, l'égalité (4) est triviale. Dans ce cas en effet, la variable $q^n$ est une constante nulle:

$$\{0,\ 0,\ 0,\ \ldots\}.$$

EXEMPLE 9. Représentons un nombre $a>0$ par sa forme décimale illimitée:

$$a=a_0,\ a_1a_2a_3\ \ldots$$

Pour sa $n$-troncature

$$a^{(n)}=a_0,\ a_1\ \ldots\ a_n$$

on a

$$\lim_{n\to\infty} a^{(n)}=a. \tag{5}$$

En effet

$$|a-a^{(n)}|=\underbrace{0,\ 0\ \ldots\ 0}_{n \text{ fois}}a_{n+1}a_{n+2}\ \ldots\leqslant 10^{-n}.$$

Si l'on se donne $\varepsilon>0$, on peut toujours exhiber un $n_0$ tel que

$$10^{-n}<\varepsilon,\quad \forall n>n_0$$

(voir exemple précédent dans lequel il faut poser $q=10^{-1}$). Donc

$$|a-a^{(n)}|<\varepsilon,\quad \forall n>n_0,$$

ce qui prouve (5).

REMARQUE 2. Les troncatures $a^{(n)}$ $(n=1,\ 2,\ \ldots)$ sont des nombres rationnels. De (5) il s'ensuit que *tout nombre réel* est la limite d'une suite de nombres rationnels.

Donc, tout nombre irrationnel peut être approché par un nombre rationnel avec n'importe quelle précision.

Cette propriété nous fait dire de l'ensemble **Q** des nombres rationnels qu'il est *partout dense* dans l'ensemble **R** des réels.

L'inégalité

$$| x_n - a | < \varepsilon$$

est équivalente aux deux inégalités

$$-\varepsilon < x_n - a < \varepsilon \quad \text{ou} \quad a - \varepsilon < x_n < a + \varepsilon.$$

Elle exprime encore le fait que le point $x_n$ appartient à l'ε-voisinage du point $a$:

$$x_n \in ]a - \varepsilon, a + \varepsilon[ \quad \text{(cf. § 10, chap. 1).}$$

On peut encore définir la limite d'une suite de la manière suivante : on dit qu'un nombre $a$ est la limite d'une variable $x_n$ si pour

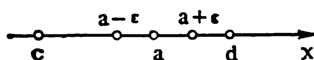


Fig. 7

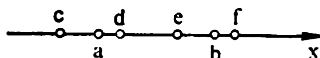


Fig. 8

tout $\varepsilon > 0$ on peut exhiber un nombre $n_0$ tel que tous les points $x_n$ d'indice $n > n_0$ soient contenus dans un ε-voisinage du point $a$:

$$x_n \in ]a - \varepsilon, a + \varepsilon[ \quad (n > n_0).$$

Il est évident que, quel que soit le voisinage $]c, d[$ du point $a$, il existe un $\varepsilon > 0$ tel que l'intervalle $]a - \varepsilon, a + \varepsilon[$ soit contenu dans $]c, d[$, c'est-à-dire que $]a - \varepsilon, a + \varepsilon[ \subset ]c, d[$ (fig. 7).

Donc, le fait que $x_n \to a$ s'exprime encore ainsi : tous les points $x_n$ à partir d'un certain indice $n$ sont contenus dans un voisinage $]c, d[$ donné à l'avance du point $a$, c'est-à-dire qu'il existe un nombre $n_0$ tel que $x_n \in ]c, d[$ $(n > n_0)$. S'agissant des points $x_n$ d'indice $n \leqslant n_0$, ils peuvent appartenir ou non à $]c, d[$. Donc, les points $x_n$ non contenus dans $]c, d[$ sont en nombre fini.

Par ailleurs, si l'on sait qu'en dehors de $]c, d[$ il n'existe qu'un nombre fini de points $x_{n_1}, x_{n_2}, \ldots, x_{n_s}$, alors en posant

$$k = \max\{n_1, n_2, \ldots, n_s\}$$

on peut dire que les points $x_n$ d'indice $n > k$ appartiennent à l'intervalle $]c, d[$. Donc, on peut formuler la notion de limite de la manière suivante : on dit qu'une variable $x_n$ tend vers $a$ si l'*ensemble des points $x_n$ situés en dehors de tout voisinage de $a$ est fini ou vide.*

Exemple 10. La suite

$$\{(-1)^{n+1}\} = \{1, -1, 1, -1, \ldots\} \tag{6}$$

ne tend vers aucune limite.

Supposons que cette suite tend vers $a$. Considérons le voisinage suivant:

$$\left]a-\frac{1}{3}, a+\frac{1}{3}\right[.$$

Il est de longueur 2/3. Il est évident que ce voisinage ne peut contenir à la fois le point 1 et le point —1, car la distance de ces points est supérieure à 2/3 ($2 > 2/3$). Supposons pour fixer les idées que le point 1 n'appartient pas au voisinage considéré. Or, $x_n = 1$ pour $n = 1, 3, 5, \ldots$, c'est-à-dire qu'un nombre infini d'éléments de cette suite sont extérieurs au voisinage considéré.

Donc, le point $a$ ne peut être limite de la suite et, comme il est arbitraire, la suite (6) n'a pas de limite.

THÉORÈME 1. *Si une variable $x_n$ admet une limite, celle-ci est unique.*

DÉMONSTRATION. Supposons par absurde que $x_n$ admet deux limites distinctes $a$ et $b$. Considérons deux intervalles disjoints $]c, d[$ et $]e, f[$ contenant respectivement les points $a$ et $b$ (fig. 8). Comme $x_n \to a$, l'intervalle $]c, d[$ contient tous les éléments $x_n$ à l'exception d'un nombre fini d'entre eux, donc l'intervalle $]e, f[$ ne peut pas contenir une infinité d'éléments $x_n$, et $b$ ne peut être limite de $x_n$. Cette contradiction prouve le théorème.

THÉORÈME 2. *Toute suite convergente est bornée.*

DÉMONSTRATION. Soit $\lim x_n = a$. Prenons $\varepsilon = 1$ et choisissons un entier naturel $n_0 = n_0(1)$ tel que

$$1 > |x_n - a| \quad (n > n_0);$$

alors $1 > |x_n| - |a|$ et l'on a

$$1 + |a| > |x_n|$$

pour tous les $n > n_0$. Soit

$$M = \max\{1 + |a|, |x_1|, |x_2|, \ldots, |x_{n_0}|\}.$$

Il est évident que

$$M \geqslant |x_n|, \quad \forall n \in \mathbf{N}.$$

Ce que nous voulions.

REMARQUE 3. Le fait qu'une suite est bornée est une condition nécessaire mais non suffisante de convergence ainsi qu'il ressort de l'exemple 10.

THÉORÈME 3. *Si une variable $x_n$ admet une limite $a$ non nulle, alors on peut exhiber un $n_0$ tel que*

$$|x_n| > |a|/2 \text{ pour } n > n_0.$$

*Bien plus, pour $n > n_0$, $x_n > a/2$ si $a > 0$, et $x_n < a/2$ si $a < 0$. Donc, $x_n$ est du signe de $a$ à partir d'un certain indice.*

DÉMONSTRATION. Supposons que $x_n \to a$. Pour $\varepsilon = |a|/2$ il existe un indice $n_0$ tel que

$$|a|/2 > |a - x_n| \geqslant |a| - |x_n| \quad \text{pour} \quad n > n_0,$$

d'où $|x_n| > |a| - \frac{|a|}{2} = \frac{|a|}{2}$, ce qui prouve la première proposition du théorème. D'autre part, l'inégalité $|a|/2 > |a - x_n|$ est équivalente aux inégalités suivantes:

$$a - \frac{|a|}{2} < x_n < a + \frac{|a|}{2} \quad (n > n_0).$$

Si $a > 0$,

$$x_n > \frac{a}{2} = a - \frac{|a|}{2} \quad (n > n_0),$$

si $a < 0$,

$$x_n < a + \frac{|a|}{2} = a - \frac{a}{2} = \frac{a}{2} \quad (n > n_0),$$

ce qui prouve la deuxième proposition du théorème.

THÉORÈME 4. *Si $x_n \to a$, $y_n \to b$ et $x_n \leqslant y_n$ pour tous les $n = 1, 2, \ldots$, alors $a \leqslant b$.*

DÉMONSTRATION. Supposons que $b < a$. Prenons $0 < \varepsilon < (a - b)/2$ et choisissons des nombres $N_1$ et $N_2$ tels que

$$a - \varepsilon < x_n \quad (n > N_1), \quad y_n < b + \varepsilon \quad (n > N_2),$$

ce qui est possible puisque $x_n \to a$ et $y_n \to b$.

Si $n_0 = \max\{N_1, N_2\}$, il est alors évident que $y_n < b + \varepsilon < a - \varepsilon < x_n$ $(n > n_0)$, ce qui est absurde puisque par hypothèse $x_n \leqslant y_n$ pour tous les $n$.

COROLLAIRE. *Si les termes d'une suite convergente $\{x_n\}$ appartiennent à $[a, b,]$ il en est de même de la limite de cette suite.*

DÉMONSTRATION. En effet, $a \leqslant x_n \leqslant b$. Si $\lim x_n = c$, alors le théorème 4 nous dit que $a \leqslant c \leqslant b$, c.q.f.d.

REMARQUE 4. S'agissant de l'intervalle $]a, b[$, on peut seulement affirmer que si $x_n \in ]a, b[$, alors $\lim x_n = c \in [a, b]$.

Donc, à la limite les inégalités se conservent ou se transforment en égalités. Par exemple, $x_n = 1/(n + 1) \in ]0, 1[$, mais $c = 0 \in [0, 1]$.

THÉORÈME 5. *Si des variables $x_n$ et $y_n$ convergent vers un même nombre $a$ et si $x_n \leqslant z_n \leqslant y_n$ $(n = 1, 2, \ldots)$, alors il en est de même de la variable $z_n$.*

DÉMONSTRATION. En se donnant $\varepsilon > 0$ on peut exhiber des nombres $N_1$ et $N_2$ tels que

$$a - \varepsilon < x_n \quad (n > N_1), \quad y_n < a + \varepsilon \quad (n > N_2),$$

d'où pour $n > n_0 = \max\{N_1, N_2\}$

$$a - \varepsilon < x_n \leqslant z_n \leqslant y_n < a + \varepsilon$$

et

$$| z_n - a | < \varepsilon \quad (n > n_0),$$

c.q.f.d.

THÉORÈME 6. *Si* $x_n \to a$, *alors* $| x_n | \to | a |$.

La démonstration résulte de l'inégalité $|| x_n | - | a || \leqslant \leqslant | x_n - a |$.

## § 2. Opérations arithmétiques sur les suites convergentes

Soient $\{x_n\}$ et $\{y_n\}$ des suites. On appelle *somme*, *différence*, *produit* et *quotient* des suites $\{x_n\}$ et $\{y_n\}$ respectivement les suites $\{x_n + y_n\}$, $\{x_n - y_n\}$, $\{x_n y_n\}$, $\{x_n/y_n\}$. Dans le cas du quotient on admet que $y_n \neq 0$ pour tous les $n = 1, 2, \ldots$

Si $x_n = c$ pour $n = 1, 2, \ldots$, on écrit $\{c \pm y_n\}$, $\{cy_n\}$, $\{c/y_n\}$ au lieu de $\{x_n \pm y_n\}$, $\{x_n y_n\}$, $\{x_n/y_n\}$.

On a les relations suivantes

$$\lim (x_n \pm y_n) = \lim x_n \pm \lim y_n, \tag{1}$$

$$\lim (x_n y_n) = \lim x_n \cdot \lim y_n, \tag{2}$$

$$\lim \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n} \quad \text{si } \lim y_n \neq 0. \tag{3}$$

Ces expressions traduisent le fait que *si les suites* $\{x_n\}$ *et* $\{y_n\}$ *admettent des limites finies, il en est de même de leur somme, de leur différence, de leur produit et de leur quotient* (*avec la réserve indiquée*) *et l'on a les relations* (1) *à* (3).

DÉMONSTRATION. Supposons que $x_n \to a$ et $y_n \to b$. Considérons $\varepsilon > 0$ et choisissons $n_0$ tel que

$$| x_n - a | < \varepsilon/2, \quad | y_n - b | < \varepsilon/2 \quad (n > n_0).$$

Alors

$$| (x_n \pm y_n) - (a \pm b) | \leqslant | x_n - a | + | y_n - b | < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (n > n_0),$$

ce qui prouve (1).

Pour démontrer (2) on remarquera que

$$| x_n y_n - ab | = | x_n y_n - a y_n + a y_n - ab | \leqslant | x_n y_n - a y_n | + \\ + | a y_n - ab | = | y_n | \, | x_n - a | + | a | \, | y_n - b |. \tag{4}$$

Comme $y_n$ admet une limite, le théorème 2 du paragraphe précédent affirme l'existence d'un nombre strictement positif $M$ tel que

$$|y_n| \leqslant M \quad (n = 1, 2, \ldots), \tag{5}$$

$$|a| \leqslant M. \tag{6}$$

Choisissons un nombre $n_0$ tel que

$$|x_n - a| < \frac{\varepsilon}{2M}, \quad |y_n - b| < \frac{\varepsilon}{2M} \quad (n > n_0). \tag{7}$$

Alors de (4) à (7) il s'ensuit que

$$|x_n y_n - ab| < \frac{M\varepsilon}{2M} + \frac{M\varepsilon}{2M} = \varepsilon \quad (n > n_0).$$

Ceci prouve (2).

Supposons maintenant que $x_n \to a$, $y_n \to b$ et que de plus $b \neq 0$. On a alors

$$\left|\frac{x_n}{y_n} - \frac{a}{b}\right| = \left|\frac{x_n b - y_n a}{y_n b}\right| = \frac{|(x_n - a)b + (b - y_n)a|}{|y_n|\,|b|} \leqslant$$
$$\leqslant \frac{|x_n - a|}{|y_n|} + \frac{|b - y_n|\,|a|}{|y_n|\,|b|}. \tag{8}$$

Le théorème 3 du paragraphe précédent nous dit que

$$|y_n| > |b|/2 \quad (n > N_1) \tag{9}$$

pour $N_1$ assez grand. Donnons-nous $\varepsilon > 0$ et choisissons $N_2$ et $N_3$ tels que

$$|x_n - a| < \varepsilon\,|b|/4 \quad (n > N_2), \tag{10}$$

$$|a|\,|y_n - b| < \varepsilon b^2/4 \quad (n > N_3). \tag{11}$$

En posant $n_0 = \max\{N_1, N_2, N_3\}$ on aura en vertu de (8) à (11)

$$\left|\frac{x_n}{y_n} - \frac{a}{b}\right| < \frac{\varepsilon \cdot |b|}{4} \cdot \frac{2}{|b|} + \frac{\varepsilon}{2} = \varepsilon \quad (n > n_0),$$

ce qui prouve (3).

Signalons que les limites des premiers membres de (1), (2) et (3) peuvent exister sans que pour autant existent séparément les limites de $x_n$ et de $y_n$. Par exemple, $x_n = (-1)^n$ et $y_n = (-1)^{n+1}$ n'admettent pas de limite. alors que $\lim (x_n + y_n) = 0$, $\lim x_n y_n = -1$.

Les théorèmes des limites d'une somme, d'une différence, d'un produit et d'un quotient nous permettent d'apprendre si la limite d'une suite existe et de la calculer si cette suite est le résultat d'un nombre fini d'opérations arithmétiques sur d'autres suites dont on connaît les limites.

Cependant de nombreux cas échappent aux applications des théorèmes indiqués et seule l'intuition nous guide au résultat escompté.

EXEMPLE. Soit $x_n = 1 + q + \ldots + q^n$, $|q| < 1$. Montrer que

$$\lim_{n\to\infty} x_n = \frac{1}{1-q}.$$

On a

$$x_n = \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q} - \frac{q^{n+1}}{1-q}.$$

Comme $\lim q^{n+1} = 0$ pour $|q| < 1$, les formules (1) et (2) nous donnent

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{1}{1-q} - \frac{1}{1-q} \lim_{n\to\infty} q^{n+1} = \frac{1}{1-q} - \frac{1}{1-q}\cdot 0 = \frac{1}{1-q}.$$

Dans la suite par

$$1+q+\ldots+q^n+\ldots \equiv \sum_{n=0}^{\infty} q^n$$

on comprendra $\lim\limits_{n\to\infty} \sum\limits_{k=0}^{n} q^k$. Donc

$$\sum_{n=0}^{\infty} q^n = \lim_{n\to\infty} x_n = \frac{1}{1-q} \quad (|q| < 1).$$

## § 3. Infiniment petit et infiniment grand

On appelle *infiniment petit* une variable $\alpha_n$ dont la limite est zéro.

Donc, une variable $\alpha_n$ est un infiniment petit si pour tout $\varepsilon > 0$ il existe un $n_0$ tel que $|\alpha_n| < \varepsilon$ pour $n > n_0$.

Il est aisé de voir que *pour qu'une variable $x_n$ admette une limite $a$ il est nécessaire et suffisant que $x_n = a + \alpha_n$, où $\alpha_n$ est un infiniment petit.*

On dit qu'une variable $\beta_n$ est un *infiniment grand* si pour tout $M > 0$ on peut exhiber un $n_0$ tel que $|\beta_n| > M$ pour $n > n_0$. On note

$$\lim \beta_n = \infty \quad \text{ou} \quad \beta_n \to \infty \tag{1}$$

et on dit que $\beta_n$ *tend vers l'infini.*

Si un infiniment grand $\beta_n$ ne prend à partir d'un indice $n_0$ que des valeurs positives ou des valeurs négatives, alors on écrit

$$\lim \beta_n = +\infty \quad \text{ou} \quad \beta_n \to +\infty, \tag{2}$$

$$\lim \beta_n = -\infty \quad \text{ou} \quad \beta_n \to -\infty. \tag{3}$$

Donc, de (2) de même que de (3) résulte (1). L'exemple de la variable $\{(-1)^n n\}$ montre qu'on peut avoir (1) mais ni (2) ni (3).

Signalons les propriétés évidentes suivantes:

1. *Si une variable* $x_n$ *est bornée et* $y_n$ *un infiniment grand, alors* $x_n/y_n \to 0$.

2. *Si la valeur absolue de* $x_n$ *est minorée par un nombre strictement positif et* $y_n$ *est un infiniment petit non nul, alors* $x_n/y_n \to \infty$.

Prouvons la deuxième propriété seulement. Par hypothèse, il existe un nombre $a > 0$ tel que $|x_n| > a$ $(n = 1, 2, \ldots)$, et pour tout $\varepsilon > 0$ un nombre $n_0$ tel que

$$|y_n| < \varepsilon \quad (n > n_0). \tag{4}$$

Alors

$$\left|\frac{x_n}{y_n}\right| > \frac{a}{\varepsilon} = M \quad (n > n_0).$$

Donnons-nous un nombre arbitraire $M > 0$ et choisissons $\varepsilon$ tel que $M = a/\varepsilon$, et $n_0(\varepsilon)$, tel que l'on ait (4). Alors $|x_n/y_n| > M$ $(n > n_0)$, c.q.f.d.

Des deux propriétés ci-dessus on déduit les corollaires suivants:

$$\lim_{y_n \to \infty} \frac{c}{y_n} = 0, \quad \lim_{y_n \to 0} \frac{c}{y_n} = \infty \quad (c \neq 0).$$

Signalons qu'une suite $\{x_n\}$ non bornée ne tend pas forcément vers l'infini. Citons à titre d'exemple la suite

$$\{n^{(-1)^n}\} = \left\{1, 2, \frac{1}{3}, 4, \ldots\right\}$$

qui n'est pas bornée et ne tend pas vers l'infini, car elle contient des termes aussi petits que l'on veut, affectés d'indices (impairs) aussi grands que l'on veut.

Remarque. De toutes les constantes, seul zéro est un infiniment petit. Si l'on sait qu'une quantité est constante et que sa valeur absolue est inférieure à tout nombre $\varepsilon > 0$, alors elle est nulle.

Théorème. *Le produit d'un infiniment petit par une quantité bornée est un infiniment petit, autrement dit si* $\lim x_n = 0$ *et* $|y_n| \leqslant M$, $\forall n \in \mathbb{N}$, *alors* $\lim x_n y_n = 0$.

En effet, donnons-nous $\varepsilon > 0$ et choisissons $n_0$ tel que

$$|x_n| < \varepsilon/M, \quad \forall n > n_0.$$

Alors

$$|x_n y_n - 0| = |x_n|\,|y_n| \leqslant \frac{\varepsilon}{M} M = \varepsilon, \quad \forall n > n_0,$$

c.q.f.d.

## § 4. Formes indéterminées

1. Supposons que $\lim x_n = \lim y_n = 0$ $(y_n \neq 0)$.

Considérons la suite $\{x_n/y_n\}$. On ne peut rien dire à priori sur la limite d'une telle suite, ainsi que le montrent les exemples suivants:

si $x_n = \frac{1}{n}$, $y_n = \frac{1}{n^2}$, alors $\frac{x_n}{y_n} = n \to +\infty$ pour $n \to \infty$;

si $x_n = \frac{1}{n^2}$, $y_n = \frac{1}{n}$, alors $\frac{x_n}{y_n} = \frac{1}{n} \to 0$ pour $n \to \infty$;

si $x_n = \frac{a}{n}$, $y_n = \frac{1}{n}$, alors $\frac{x_n}{y_n} = a \to a$ pour $n \to \infty$;

si $x_n = \frac{(-1)^n}{n}$, $y_n = \frac{1}{n}$, alors $\frac{x_n}{y_n} = (-1)^n$ et cette suite n'a pas de limite.

Donc, pour calculer la limite de $\{x_n/y_n\}$ la seule connaissance du fait que $x_n \to 0$ et $y_n \to 0$ ne suffit pas. Il faut encore être fixé sur le caractère de variation de $x_n$ et de $y_n$. On aura besoin de procédés spéciaux pour trouver cette limite dans chaque cas concret.

On dit que l'expression $x_n/y_n$ pour $x_n \to 0$ et $y_n \to 0$ est une forme *indéterminée de type* $\left(\frac{0}{0}\right)$.

2. Si $x_n \to \infty$ et $y_n \to \infty$, l'expression $x_n/y_n$ donne lieu à une *indétermination de la forme* $\left(\frac{\infty}{\infty}\right)$.

3. Si $x_n \to 0$ et $y_n \to \infty$, l'expression $x_n y_n$ donne lieu à une *indétermination de la forme* $(0 \cdot \infty)$.

4. Si $x_n \to +\infty$ et $y_n \to -\infty$, l'expression $x_n + y_n$ donne lieu à une *indétermination de la forme* $(\infty - \infty)$.

Lever l'indétermination c'est trouver la limite (si elle existe) de l'expression considérée, ce qui n'est pas toujours facile.

EXEMPLE 1. Si

$$x_n = a_m n^m + \ldots + a_1 n + a_0,$$

$$y_n = b_l n^l + \ldots + b_1 n + b_0 \quad (a_m \neq 0, \quad b_l \neq 0),$$

le quotient $x_n/y_n$ donne lieu à une indétermination de la forme $\left(\frac{\infty}{\infty}\right)$ pour $n \to \infty$. Levons cette indétermination.

a) Si $l = m$, une division du numérateur et du dénominateur par $n^m$ nous donne

$$\frac{x_n}{y_n} = \frac{a_m + \frac{a_{m-1}}{n} + \ldots + \frac{a_0}{n^m}}{b_m + \frac{b_{m-1}}{n} + \ldots + \frac{b_0}{n^m}}.$$

On voit maintenant que $\lim_{n\to\infty} (x_n/y_n) = a_m/b_m$, c'est-à-dire est égale au quotient des coefficients supérieurs de $x_n$ et de $y_n$.

b) On démontre de façon analogue que $\lim_{n\to\infty} (x_n/y_n) = \infty$ pour $m > l$, et $\lim_{n\to\infty} (x_n/y_n) = 0$ pour $m < l$.

EXEMPLE 2. Si $x_n = \sqrt{n+1}$, $y_n = \sqrt{n}$, le calcul de la limite de l'expression $x_n - y_n$ pour $n \to \infty$ nous place dans le cas d'une indétermination de la forme $(\infty - \infty)$.

Levons cette indétermination:

$$x_n - y_n = \sqrt{n+1} - \sqrt{n} = \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n+1}+\sqrt{n}} (\sqrt{n+1}+\sqrt{n}) =$$

$$= \frac{1}{\sqrt{n+1}+\sqrt{n}}.$$

Maintenant $\lim_{n\to\infty} (\sqrt{n+1} - \sqrt{n}) = 0$.

## § 5. Suites monotones

DÉFINITION. On dit qu'une suite $\{x_n\}$ est *croissante* (resp. *décroissante*) si pour tout $n \in \mathbf{N}$ on a

$$x_n \leqslant x_{n+1} \quad (\text{resp. } x_n \geqslant x_{n+1}).$$

Si $x_n < x_{n+1}$ (resp. $x_n > x_{n+1}$), on dit que la suite $\{x_n\}$ est *strictement croissante* (resp. *strictement décroissante*). Les suites croissantes, décroissantes (resp. strictement croissantes, strictement décroissantes) sont dites *monotones* (resp. *strictement monotones*).

Il est immédiat de voir que la suite croissante $x_1 \leqslant x_2 \leqslant \ldots \leqslant \leqslant x_n \leqslant x_{n+1} \leqslant \ldots$ et la suite décroissante $x_1 \geqslant x_2 \geqslant \ldots \geqslant \geqslant x_n \geqslant x_{n+1} \geqslant \ldots$ sont respectivement minorée et majorée.

EXEMPLE 1.

1) $\{1, 1, 1/2, 1/2, \ldots 1/n, 1/n, \ldots\}$ est une suite décroissante.

2) $\{n^2\}$ est une suite strictement croissante.

On démontre plus bas un théorème important qui dit que toute suite monotone minorée ou majorée est convergente. Au § 7, chap. 1, ce théorème figurait comme une propriété fondamentale (la propriété V) de l'ensemble $\mathbf{R}$ des réels.

THÉORÈME. *Si une suite de nombres réels*

$$a_1, a_2, a_3, \ldots \tag{1}$$

*est croissante (resp. décroissante) et majorée (resp. minorée) par un nombre $M$ (resp. $m$), alors cette suite converge vers un nombre $a$ infé-*

*rieur à M (resp. supérieur à m)*:

$$\lim_{n\to\infty} a_n = a \leqslant M \tag{2}$$

(*resp.* $\lim_{n\to\infty} a_n = a \geqslant m$).

DÉMONSTRATION. Supposons que la suite (1) est croissante et que $a_1 > 0$, donc que tous les $a_n > 0$ ($n = 1, 2, \ldots$). Représentons chaque terme de la suite par une forme décimale illimitée

$$a_n = a_{n0}, \; a_{n1}a_{n2}a_{n3} \ldots \quad (n = 1, 2, \ldots). \tag{3}$$

La suite $\{a_n\}$ étant majorée par $M$ est croissante, le lemme 2 du § 6, chap. 1, nous dit que les formes décimales (3) se stabilisent à une forme correspondant à un nombre $a \leqslant M$:

$$a_n \rightrightarrows a = \gamma_0, \; \gamma_1\gamma_2 \ldots,$$

donc $a$ est la limite de $a_n$:

$$\lim_{n\to\infty} a_n = a.$$

En effet, pour tout $\varepsilon$ il existe un entier naturel $m$ tel que $10^{-m} < \varepsilon$. Comme $a_n$ se stabilise à $a$, on a

$$a_n = \gamma_0, \; \gamma_1\gamma_2 \ldots \gamma_m a_{n,m+1} a_{n,m+2} \ldots$$

pour tous les $n > n_0$, où $n_0$ est assez grand. Mais alors

$$|a - a_n| = a - a_n \leqslant 0, \underbrace{0 \ldots 0}_{m \text{ fois}} \gamma_{m+1}\gamma_{m+2} \ldots \leqslant 10^{-m} < \varepsilon,$$

c'est-à-dire que $a_n \to a$ pour $n \to \infty$.

Si $a_1 \leqslant 0$, on lui ajoute un nombre $c$ tel que $a_1 + c > 0$ et on pose $b_n = a_n + c$ ($n = 1, 2, \ldots$).

La suite $\{b_n\}$ est croissante, majorée par le nombre $M + c$ et tous ses termes sont strictement positifs. Donc, d'après le théorème que nous venons de prouver l'existence de $\lim_{n\to\infty} b_n = b \leqslant M + c$ entraîne celle de $\lim_{n\to\infty} a_n = \lim_{n\to\infty} (b_n - c) = b - c \leqslant M$, ce qui prouve le théorème pour toute suite croissante.

Si maintenant la suite $\{a_n\}$ est décroissante et minorée par un nombre $m$, la suite de nombres $\{-a_n\}$ est croissante et majorée par le nombre $-m$, et le théorème affirme l'existence de $\lim_{n\to\infty} (-a_n) = -a \leqslant -m$. Donc, il existe également $\lim_{n\to\infty} a_n = -\lim_{n\to\infty} (-a_n) = -(-a) = a \geqslant m$. Ceci prouve le théorème.

REMARQUE. Si une suite de nombres réels $\{a_n\}$ converge, les formes décimales de ses termes ne se stabilisent pas nécessairement. Si par exemple

$$a_{2k} = 1,\underbrace{0 \ldots 0}_{k \text{ fois}}11 \ldots \quad (k = 1, 2, \ldots),$$

$$a_{2k+1} = 0,\underbrace{9 \ldots 9}_{k \text{ fois}}11 \ldots \quad (k = 1, 2, \ldots),$$

la suite $\{a_n\}$ converge vers 1 ($a_n \to 1$), mais il est aisé de voir que cette suite ne se stabilise pas.

Exemple 2. Voici une autre démonstration de l'égalité (cf. exemple 8, § 1)

$$\lim_{n\to\infty} q^n = 0, \quad |q| < 1. \tag{4}$$

Supposons d'abord que $q \geqslant 0$. La variable $q^n$ ($n = 1, 2, \ldots$) est décroissante et minorée par 0. Le théorème affirme l'existence d'un nombre $A \geqslant 0$ qui est la limite de $q^n$:

$$\lim_{n\to\infty} q^n = A.$$

On a par ailleurs

$$A = \lim_{n\to\infty} q^{n+1} = q \lim_{n\to\infty} q^n = qA,$$

d'où $A(1-q) = 0$, et $A = 0$, car $q < 1$.

Si maintenant $q < 0$, alors d'après ce qui vient d'être démontré $|q^n| = |q|^n \to 0$, $n \to \infty$.

Ceci achève la démonstration de (4).

Cette démonstration est plus élégante que celle donnée au § 1, exemple 8, mais elle ne permet pas de juger de la vitesse de convergence de $q^n$ vers 0, c'est-à-dire qu'elle n'indique pas le nombre $n_0 = n_0(\varepsilon)$ à partir duquel $|q^n| < \varepsilon$.

Exemple 3. On a l'égalité

$$\lim_{n\to\infty} \frac{a^n}{n!} = 0, \tag{5}$$

où $a$ est un nombre quelconque.

Pour $|a| \leqslant 1$, cette relation est évidente. Supposons que $a > 1$. Posons

$$u_n = \frac{a^n}{n!}.$$

Alors

$$\frac{u_{n+1}}{u_n} = \frac{a}{n+1} \to 0 \quad (n \to \infty).$$

D'où il suit que $u_{n+1} < u_n$, $\forall\, n > n_0$, où $n_0$ est assez grand.

Donc, la variable $u_n$ décroît pour $n > n_0$. De plus, elle est minorée par 0. Donc, il existe

$$\lim_{n\to\infty} u_n = A \geqslant 0.$$

On a aussi

$$A = \lim_{n\to\infty} u_{n+1} = \lim_{n\to\infty}\left(u_n \frac{a}{n+1}\right) = A \lim_{n\to\infty} \frac{a}{n+1} = A\cdot 0 = 0,$$

ce qui prouve (5) pour tout $a \geqslant 0$. Or, elle est valable pour tout $a < 0$, car

$$\left|\frac{a^n}{n!}\right| = \frac{|a|^n}{n!} \to 0 \text{ pour } n \to \infty.$$

## § 6. Le nombre $e$

Considérons la suite

$$\{x_n\} = \left\{\left(1+\frac{1}{n}\right)^n\right\}.$$

Montrons que cette suite est strictement croissante et majorée. En vertu de la formule du binôme de Newton

$$(a+b)^n = \sum_{k=0}^{n} \frac{n(n-1)\ldots(n-k+1)}{k!} a^{n-k} b^k$$

on a

$$x_n = \left(1+\frac{1}{n}\right)^n = 1 + n\frac{1}{n} + \ldots$$

$$\ldots + \frac{n(n-1)\ldots(n-k+1)}{k!}\frac{1}{n^k} + \ldots + \frac{n(n-1)\ldots(n-n+1)}{n!}\cdot\frac{1}{n^n} =$$

$$= 2 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \ldots + \frac{1}{k!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{k-1}{n}\right) + \ldots$$

$$\ldots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{n-1}{n}\right). \quad (1)$$

On constate que $x_n \geqslant 2$ pour tout $n$. Montrons que la suite $\{x_n\}$ est majorée. De (1) il vient

$$x_n \leqslant 2 + \frac{1}{2!} + \ldots + \frac{1}{n!} \leqslant 2 + \frac{1}{2} + \ldots + \frac{1}{2^{k-1}} + \ldots + \frac{1}{2^{n-1}} \leqslant$$

$$\leqslant 1 + \left(1 + \frac{1}{2} + \ldots + \frac{1}{2^n} + \ldots\right) = 1 + \frac{1}{1-\frac{1}{2}} = 3.$$

Montrons que la suite $\{x_n\}$ est strictement croissante. Par analogie avec (1) on a

$$x_{n+1} = 2 + \frac{1}{2!}\left(1-\frac{1}{n+1}\right) + \ldots$$

$$\ldots + \frac{1}{k!}\left(1-\frac{1}{n+1}\right)\ldots\left(1-\frac{k-1}{n+1}\right) + \ldots$$

$$\ldots + \frac{1}{n!}\left(1-\frac{1}{n+1}\right)\ldots\left(1-\frac{n-1}{n+1}\right) +$$

$$+ \frac{1}{(n+1)!}\left(1-\frac{1}{n+1}\right)\ldots\left(1-\frac{n}{n+1}\right). \quad (2)$$

En comparant (1) et (2) on voit que $x_n < x_{n+1}$, $\forall n \in \mathbf{N}$ (chaque terme de (2) est strictement supérieur au terme respectif de (1) et en outre (2) contient un terme positif de plus que (1)). Le théorème du § 5 nous dit que la suite $\{x_n\}$ converge. Désignons cette limite par $e$ comme l'a proposé Euler:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e.$$

De ce qui précède il est clair que $2 < e < 3$. Le nombre $e$ est un nombre transcendant dont les premières décimales sont

$$2{,}718\ 281\ 828\ 459\ \ldots$$

Au § 16, chap. 4, on prouvera une formule d'où il résulte que

$$e=\sum_{k=0}^{n}\frac{1}{k!}+\frac{\theta}{n!}\quad (n>2), \tag{3}$$

où $\theta$ est un nombre dépendant de $n$, tel que $0 < \theta < 1$. On démontre facilement à l'aide de cette formule que $e$ est un nombre irrationnel. Supposons que $e = p/q$, où $p$ et $q$ sont des entiers naturels. En posant $n = q$ dans (3), on obtient

$$\frac{p}{q}=\sum_{k=0}^{q}\frac{1}{k!}+\frac{\theta}{q!}.$$

En multipliant par $q!$ on trouve

$$p\,(q-1)!-l=\theta, \tag{4}$$

où $l = q!\sum_{k=0}^{q}\frac{1}{k!}$ est un entier naturel. Le premier membre de (4) est un entier, tandis que le second, $\theta$, est une fraction propre.

## § 7. Principe des segments emboîtés

THÉORÈME DES SEGMENTS EMBOÎTÉS *Soit donnée une suite d'intervalles fermés*

$$\sigma_n = [a_n, b_n] \quad (n = 1, 2, \ldots)$$

*emboîtés, c'est-à-dire tels que* $\sigma_{n+1} \subset \sigma_n$ $(n = 1, 2, \ldots)$, *dont les longueurs tendent vers* 0:

$$d_n = b_n - a_n \to 0 \quad (n \to \infty).$$

*Il existe alors un point unique c intérieur à tous ces intervalles* $(c \in \sigma_n,\ n = 1, 2, \ldots)$.

DÉMONSTRATION. Il est évident que

$$a_1 \leqslant a_2 \leqslant a_3 \leqslant \ldots \leqslant b_m$$

pour tout $m$. La suite de nombres $a_n$ est croissante et majorée par $b_m$ pour tout $m$, donc en vertu du théorème du § 5 il existe un nombre $c$ vers lequel elle converge ($\lim a_n = c$). De plus $a_n \leqslant c \leqslant b_m$. Comme

les nombres $n$ et $m$ sont arbitraires dans ces inégalités, on a en particulier $a_n \leqslant c \leqslant b_n$ $(n = 1, 2, \ldots)$. Donc, $c \in \sigma_n$ pour tout $n \in \mathbf{N}$.

Le point $c$ est unique. Supposons qu'il existe un autre point $c_1 \in \sigma_n$ pour tout $n$. On a alors $a_n \leqslant c, c_1 \leqslant b_n$, d'où

$$b_n - a_n \geqslant | c - c_1 | > 0 \quad \text{pour tout } n,$$

ce qui contredit le fait que $b_n - a_n \to 0$.

Signalons que $\lim b_n = \lim [(b_n - a_n) + a_n] = c$.

Remarque. Dans le théorème ci-dessus, il est essentiel que les intervalles $[a_n, b_n]$ soient fermés, comme le montre l'exemple suivant. Les intervalles $]0, 1/n[$ $(n = 1, 2, \ldots)$ sont emboîtés et leur longueur $d_n = \frac{1}{n} - 0 = \frac{1}{n} \to 0$, mais ils ne contiennent aucun point commun.

En effet, aucun de ces intervalles ne contient de points $c \leqslant 0$. Si $c > 0$, il existe un $n$ tel que $\frac{1}{n} < c$ et $c \notin \left]0, \frac{1}{n}\right[$.

## § 8. Bornes supérieure et inférieure d'un ensemble

Considérons un ensemble $E$ de nombres réels $x$. Désignons par $M$ le plus grand élément (s'il existe) de $E$:

$$M = \max E = \max_{x \in E} x.$$

Désignons par $m$ le plus petit élément (s'il existe) de $E$:

$$m = \min E = \min_{x \in E} x.$$

Si l'ensemble $E$ est fini, c'est-à-dire est composé d'un nombre fini d'éléments

$$x_1, x_2, \ldots, x_p,$$

il contient toujours un plus petit et un plus grand.

Cependant ce n'est pas toujours le cas si $E$ est un ensemble infini.

Exemples:

1) $\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$,
2) $\mathbf{N} = \{1, 2, 3, \ldots\}$,
3) $\mathbf{N}_- = \{\ldots, -2, -1\}$,
4) $[a, b]$,
5) $[a, b[$,
6) $]a, b[$.

L'ensemble $\mathbf{Z}$ ne possède ni d'élément maximum ni d'élément minimum. Il en est de même de l'intervalle $]a, b[$, que les nombres $a$ et $b$ soient finis ou non. Pour tout nombre $c \in ]a, b[$, c'est-à-dire

tel que $a < c < b$, on peut toujours trouver des nombres $c_1$ et $c_2$ tels que $a < c_1 < c < c_2 < b$.

L'ensemble **N** ne possède pas d'élément maximum, par contre son élément minimum est 1. L'ensemble $\mathbf{N}_-$ possède un élément maximum $x = -1$ mais pas d'élément minimum.

Il est de même évident que min $[a, b] = a$, max $[a, b] = b$, min $[a, b[ = a$ et max $[a, b[$ n'existe pas.

Nous allons introduire pour l'ensemble $E$ des nombres susceptibles de remplacer éventuellement max $E$ et min $E$. Ces nombres (finis ou infinis) sont la *borne supérieure* ou *suprémum*

$$\sup E = \sup_{x \in E} x = M$$

et la *borne inférieure* ou *infimum*

$$\inf E = \inf_{x \in E} x = m.$$

Supposons que l'ensemble $E$ est majoré.

On dit qu'un nombre $M$ (fini) est la *borne supérieure* de $E$ si

1) $x \leqslant M$, $\forall x \in E$,

2) pour tout $\varepsilon > 0$ il existe un nombre $x_1 \in E$ tel que

$$M - \varepsilon < x_1 \leqslant M.$$

Supposons maintenant que l'ensemble $E$ est minoré.

On dit qu'un nombre $m$ (fini) est la *borne inférieure* de $E$ si

1) $m \leqslant x$, $\forall x \in E$,

2) pour tout $\varepsilon > 0$ il existe un nombre $x_1 \in E$ tel que

$$m \leqslant x_1 < m + \varepsilon.$$

Il est évident que si un ensemble $E$ de nombres réels possède un élément maximum (resp. minimum), c'est-à-dire si existe max $E$ (resp. min $E$), alors

$$\sup E = \max E \quad (\text{resp. } \inf E = \min E).$$

EXEMPLE 1. L'ensemble

$$E = \left\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots\right\} = \left\{\frac{n}{n+1}\right\}$$

possède un élément minimum (min $E = 1/2$) mais pas d'élément maximum car $\frac{1}{2} < \frac{2}{3} < \frac{3}{4} < \ldots$ Cet ensemble est majoré par le nombre 1 ou par tout nombre supérieur à 1.

En effet:

1) $\frac{n}{n+1} < 1$, $\forall n \in \mathbf{N}$,

2) $\forall \varepsilon > 0$, $\exists n_1 \in \mathbf{N}: 1 - \varepsilon < \frac{n_1}{n_1+1} < 1$.

Nous avons défini la borne supérieure (resp. inférieure) d'un ensemble majoré (resp. minoré).

Si l'ensemble $E$ n'est pas majoré (resp. minoré), alors $\sup E = \infty$ (resp. $\inf E = -\infty$).

Exemple 2. Pour les ensembles 1) à 6) examinés plus haut, on a

$\sup \mathbf{Z} = +\infty$, $\inf \mathbf{Z} = -\infty$,
$\sup \mathbf{N} = \infty$, $\inf \mathbf{N} = \min \mathbf{N} = 1$,
$\sup \mathbf{N}_- = \max \mathbf{N}_- = -1$, $\inf \mathbf{N}_- = -\infty$,
$\sup ]a, b[ = b$, $\inf ]a, b[ = a$,

où $a$ et $b$ peuvent être finis ou infinis.

On peut donner une définition générale de la borne supérieure (resp. inférieure) qui convient pour tout ensemble (borné ou non).

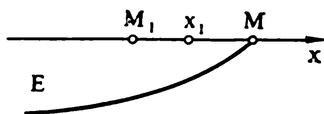



Fig. 9

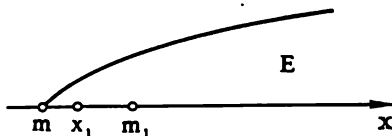



Fig. 10

On dit qu'un nombre $M$ (resp. $m$) fini ou infini est la *borne supérieure* (resp. *inférieure*) d'un ensemble $E$ (fig. 9 et 10) si

1) $x \leqslant M$ (resp. $m \leqslant x$), $\forall x \in E$;

2) pour tout nombre (fini !) $M_1 < M$ (resp. $m_1 > m$) il existe un $x_1 \in E$ tel que $M_1 < x_1 \leqslant M$ (resp. $m \leqslant x_1 < m_1$).

Dans cette formulation on ne se sert pas de la différence $M - \varepsilon$ (resp. la somme $m + \varepsilon$) qui n'a pas de sens pour $M = \infty$ (resp. $m = -\infty$).

Le théorème suivant est fondamental.

Théorème. *Si un ensemble non vide $E$ de nombres réels admet un majorant fini $K$ (resp. un minorant fini $k$), il existe alors un nombre $M \leqslant K$ (resp. $m \geqslant k$) qui est la borne supérieure (resp. inférieure) de $E$.*

Démonstration. L'ensemble $E$ n'étant pas vide, il contient au moins un point $x_0$. Considérons l'intervalle fermé $\sigma_0 = [a, b]$, où $a < x_0$, $b = K$.

Par hypothèse, il n'existe aucun point de $E$ à droite de $\sigma_0$. Partageons $\sigma_0$ en deux parties égales (deux intervalles fermés) et désignons par $\sigma_1$ la partie droite qui contient au moins un point de $E$. Si la partie de droite ne contient aucun point, on désignera celle de gauche par $\sigma_1$.

Soit maintenant $x_1$ un point de $E$ appartenant à $\sigma_1$. A droite de $\sigma_1$, il n'existe donc aucun point de $E$. Partageons $\sigma_1$ en deux intervalles fermés égaux et désignons par $\sigma_2$ l'intervalle le plus à droite contenant au moins un point de $E$.

En poursuivant ce processus on obtient une suite d'intervalles fermés emboîtés $\sigma_n = [a_n, b_n]$ $(\sigma_n \supset \sigma_{n+1})$ dont les longueurs

$$b_n - a_n = \frac{b-a}{2^n} \to 0 \quad \text{pour } n \to \infty.$$

Ceci étant, il n'existe aucun point de $E$ à droite de $\sigma_n$ quel que soit $n \in \mathbf{N}$, mais $\sigma_n$ contient au moins un point $x_n \in E$.

D'après le théorème des segments emboîtés, il existe un point unique, que nous désignerons par $M$, qui appartient à tous les intervalles $\sigma_n$.

Montrons que

$$M = \sup E. \tag{1}$$

En effet,

1) $x \leqslant M$, $\forall x \in E$, car $x \leqslant b_n$, $\forall n$, et, en passant à la limite pour $n \to \infty$, on obtient

$$x \leqslant \lim_{n\to\infty} b_n = M;$$

2) $\forall \varepsilon > 0$, $\exists x' \in E$:

$$M - \varepsilon < x' \leqslant M. \tag{2}$$

En effet, les points $x_n$ définis plus haut appartiennent à $\sigma_n$ et à $E$, c'est-à-dire que $a_n \leqslant x_n \leqslant M$, et comme $a_n \to M$ pour $n \to \infty$, alors

$$\forall \varepsilon > 0,\ \exists n_0\colon\ M - \varepsilon < a_{n_0} \leqslant x_{n_0} \leqslant M,$$

et l'on obtient (2) si l'on pose $x' = x_{n_0}$.

Le théorème se démontre de façon analogue pour la borne inférieure. On considère un intervalle $\sigma_0 = [a, b]$ contenant un point $x_0 \in E$, tel que $a = k$ et $x_0 < b$, et on le partage en deux parties égales. On désigne par $\sigma_1$ la partie la plus à gauche contenant un point $x_1 \in E$ et on poursuit ce processus.

Ce qui précède nous conduit à la proposition suivante: *tout ensemble $E$ admet une borne supérieure et une borne inférieure. Si $E$ est majoré, alors* $\sup E < \infty$, *sinon* $\sup E = \infty$. *De façon analogue, si $E$ est minoré, alors* $\inf E > -\infty$, *sinon* $\inf E = -\infty$.

EXERCICES. **1.** Soient donnés deux ensembles de nombres réels $X = \{x\}$, $Y = \{y\}$. Par ensemble $\{x + y\}$ on entendra l'ensemble de toutes les sommes des nombres $x \in X$ et $y \in Y$. Montrer que

$$\sup\{x + y\} = \sup\{x\} + \sup\{y\},$$
$$\inf\{x + y\} = \inf\{x\} + \inf\{y\}.$$

2. $\{xy\}$ étant l'ensemble de tous les produits des nombres positifs $x \in X$ et $y \in Y$, montrer que

$$\sup\{xy\} = \sup\{x\}\sup\{y\},$$
$$\inf\{xy\} = \inf\{x\}\inf\{y\} \quad (x \geqslant 0,\ y \geqslant 0).$$

3. Montrer que

$$\sup_{x\in A}(-x) = -\inf_{x\in A} x, \quad \inf_{x\in A}(-x) = -\sup_{x\in A} x.$$

## § 9. Théorème de Bolzano-Weierstrass [1]

Soit donnée une suite de nombres réels $\{x_n\}$. Choisissons parmi cette suite une infinité d'éléments d'indices $n_1 < n_2 < \ldots$ La suite $\{x_{n_k}\}$ obtenue s'appelle *sous-suite* ou *suite extraite de la suite* $\{x_n\}$. Toute suite contient une infinité de suites extraites.

Si une suite $\{x_n\}$ converge (vers un nombre fini ou infini), il est évident que chacune de ses suites extraites converge, et de plus vers le même nombre.

---

[1]) Bolzano Bernhard, mathématicien tchèque (1781-1848), Weierstrass Karl, mathématicien allemand (1815-1897).

La suite

$$\{1, -1, 1, -1, \ldots\} \qquad (1)$$

peut servir d'exemple de suite divergente. On remarque toutefois qu'elle contient une suite extraite

$$\{1, 1, 1, \ldots\}$$

qui converge vers 1. Il se pose la question de savoir si toute suite de nombres réels contient une suite extraite convergeant vers un nombre fini ou infini. Le théorème suivant répond à cette question.

THÉORÈME 1. *De toute suite de nombres réels $\{x_n\}$ on peut extraire une suite $\{x_{n_k}\}$ convergeant vers un nombre fini ou infini.*

Si la suite $\{x_n\}$ n'est pas majorée (resp. minorée), elle contient visiblement une suite extraite convergeant vers $+\infty$ (resp. $-\infty$), ce qui prouve le théorème. Si la suite est bornée, le théorème 1 se ramène au suivant.

THÉORÈME 2 (BOLZANO-WEIERSTRASS). *De toute suite bornée $\{x_n\}$ on peut extraire une suite convergente $\{x_{nk}\}$.*

DÉMONSTRATION. La suite $\{x_n\}$ étant bornée, tous ses points appartiennent à un intervalle fermé $[a, b]$ que nous désignons par $\sigma_0$. Partageons $\sigma_0$ en deux intervalles fermés égaux et désignons par $\sigma_1$ l'intervalle le plus à droite qui contient une infinité d'éléments $x_n$. Soit $x_{n_1}$ un de ces éléments. A droite de $\sigma_1$ on ne peut avoir éventuellement qu'un nombre fini de points $x_n$. Partageons l'intervalle $\sigma_1$ en deux intervalles fermés égaux et désignons par $\sigma_2$ l'intervalle le plus à droite contenant une infinité d'éléments $x_n$. Soit $x_{n_2}$, $n_2 > n_1$, un de ces éléments. S'il existe des points $x_n$ à droite de $\sigma_2$, ils ne peuvent être qu'en nombre fini.

En poursuivant ce processus on obtient une suite d'intervalles emboîtés $\sigma_k = [a_k, b_k]$, dont les longueurs $b_k - a_k \to 0$ pour $k \to \infty$, et une suite extraite de points tels que $x_{nk} \in \sigma_k$ ($n_1 < n_2 < \ldots$). Ceci étant, à droite de chaque intervalle fermé, les éléments $x_n$ sont au plus en nombre fini.

D'après le théorème des segments emboîtés, il existe un point $c$ intérieur à tous les intervalles $\sigma_k$. Il est évident que la suite extraite $\{x_{nk}\}$ tend vers $c$. Ce que nous voulions.

## § 10. Limites supérieure et inférieure

Etant donnée une suite de nombres réels $\{x_n\}$, on peut, selon le théorème 1 du § 9, en extraire une suite convergeant vers un nombre fini ou infini.

On appelle *limite supérieure d'une suite* $\{x_n\}$ (ou *d'une variable* $x_n$) un nombre $M$ (fini ou infini) possédant les deux propriétés suivantes:

1) $M$ est la limite d'une suite extraite $\{x_{n_k}\}$ de $\{x_n\}$:

$$\lim_{k\to\infty} x_{n_k} = M.$$

2) Pour toute suite extraite convergente $\{x_{n_k}\}$:

$$\lim_{k\to\infty} x_{n_k} \leqslant M.$$

On désigne la limite supérieure d'une suite $\{x_n\}$ par le symbole

$$M = \lim_{n\to\infty} \sup_{k>n} x_k.$$

Si la suite $\{x_n\}$ n'est pas majorée, il est évident que

$$\lim_{n\to\infty} \sup_{k>n} x_k = +\infty.$$

EXEMPLES.

$\lim\limits_{n\to\infty} \sup\limits_{k>n} (-1)^k = 1.$

$\lim\limits_{n\to\infty} \sup\limits_{k>n} k^{(-1)^k} = +\infty$, car cette suite n'est pas majorée.

La limite supérieure $M$ d'une suite $\{x_n\}$ majorée peut être définie de la manière suivante: un nombre $M$ est limite supérieure d'une suite $\{x_n\}$ si pour tout $\varepsilon > 0$ l'ensemble des points situés à droite de $M + \varepsilon$ est fini et celui des points situés à droite de $M - \varepsilon$, infini.

On remarquera que si une suite $\{x_n\}$ admet $M$ pour limite, alors on sait que quel que soit $\varepsilon > 0$, les inégalités $M - \varepsilon < x_n < M + \varepsilon$ sont réalisées pour tous les $x_n$ sauf pour un nombre fini d'entre eux. Donc, on a un nombre au plus fini d'éléments $x_n$ à droite de $M + \varepsilon$ et une infinité à droite de $M - \varepsilon$.

Ceci montre que $M = \lim\limits_{n\to\infty} \sup\limits_{k>n} x_k$.

Donc, si $M = \lim\limits_{n\to\infty} x_n$, alors $\lim\limits_{n\to\infty} \sup\limits_{k>n} x_k = \lim\limits_{n\to\infty} x_n = M$.

La différence entre la limite ordinaire et la limite supérieure réside dans le fait que pour la première l'ensemble des points situés à gauche de $M - \varepsilon$ est au plus fini, tandis que pour la seconde, cet ensemble est éventuellement infini.

On appelle *limite inférieure d'une suite* $\{x_n\}$ (ou *d'une variable* $x_n$) un nombre $m$ (fini ou infini) possédant les deux propriétés suivantes:

1) $m$ est la limite d'une suite extraite $\{x_{n_k}\}$ de $\{x_n\}$:

$$\lim_{k\to\infty} x_{n_k} = m.$$

2) Pour toute suite extraite convergente $\{x_{n_k}\}$:

$$\lim_{k\to\infty} x_{n_k} \geqslant m.$$

La limite inférieure de la suite $\{x_n\}$ est désignée par le symbole

$$m = \lim_{n\to\infty} \inf_{k>n} x_k.$$

Si la suite $\{x_n\}$ n'est pas minorée, il est alors évident que

$$\lim_{n\to\infty} \inf_{k>n} x_k = -\infty.$$

On peut définir la limite inférieure d'une suite minorée de la manière suivante: un nombre $m$ est limite inférieure d'une suite $\{x_n\}$ si pour tout $\varepsilon > 0$ l'ensemble des points $x_n$ situés à gauche de $m - \varepsilon$ est fini et celui des points situés à gauche de $m + \varepsilon$, infini.

Il est manifeste que

$$\lim_{n\to\infty} \inf_{k>n} x_k \leqslant \lim_{n\to\infty} \sup_{k>n} x_k. \tag{1}$$

THÉORÈME. *Pour qu'une suite $\{x_n\}$ soit convergente il est nécessaire et suffisant que* $\lim_{n\to\infty} \sup_{k>n} x_k = \lim_{n\to\infty} \inf_{k>n} x_k$. *Ceci étant,* $\lim_{n\to\infty} x_n = \lim_{n\to\infty} \inf_{k>n} x_k = \lim_{n\to\infty} \sup_{k>n} x_k$.

On remarquera que si $\lim_{n\to\infty} \sup_{k>n} x_k = -\infty$, alors en vertu de (1) $\lim_{n\to\infty} \inf_{k>n} x_k = -\infty$ et le théorème nous dit que

$$\lim_{n\to\infty} x_n = -\infty.$$

Il est évident par ailleurs que l'égalité $\lim_{n\to\infty} \inf_{k>n} x_k = +\infty$ entraîne que

$$\lim_{n\to\infty} \sup_{k>n} x_k = \lim_{n\to\infty} x_n = +\infty.$$

REMARQUE. On peut démontrer que le nombre $c$ que nous avons obtenu dans la démonstration du théorème de Bolzano-Weierstrass est la limite supérieure de $x_n$.

Ceci résulte du fait que l'ensemble des points situés à droite de chaque intervalle $\sigma_n$ est au plus fini.

D'autre part, si nous avions modifié le processus et retenu à chaque étape de la division de $\sigma_n$ en deux intervalles égaux non pas l'intervalle le plus à droite mais celui le plus à gauche qui contient une infinité de points $x_n$, nous aurions probablement obtenu un autre point $c'$ intérieur à tous les $\sigma_n$ qui aurait été la limite inférieure de $x_n$.

Si la suite $\{x_n\}$ ne possède pas de limite, alors $c' < c$, sinon ces deux processus nous conduiront à un même nombre $c = c'$.

## § 11. Critère de Cauchy de convergence d'une suite

Soit donnée une suite de nombres réels $\{x_n\}$ convergeant vers un nombre fini $a$:

$$\lim_{n\to\infty} x_n = a.$$

Il revient au même de dire que pour tout $\varepsilon > 0$ il existe un nombre $n_0 = n_0(\varepsilon)$ tel que

$$|x_n - a| < \varepsilon/2, \quad \forall\, n > n_0.$$

On a de même pour $m > n_0$:

$$|x_m - a| < \varepsilon/2, \quad \forall\, m > n_0.$$

Et

$$|x_n - x_m| = |x_n - a + a - x_m| \leqslant |x_n - a| + |x_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \forall n,\ m > n_0.$$

Nous avons obtenu la proposition suivante: *si une suite $\{x_n\}$ converge vers un nombre fini, elle remplit la condition de Cauchy* [1]:

[1]) Augustin Louis Cauchy, mathématicien français (1789-1857), le premier à avoir défini dans ses travaux les notions fondamentales de l'analyse (limite, continuité, intégrale, ...) telles qu'elles sont en usage dans les mathématiques modernes.

*pour tout* $\varepsilon > 0$ *il existe un* $n_0 = n_0(\varepsilon)$ *tel que*

$$|x_m - x_n| < \varepsilon, \quad \forall\, n,\ m > n_0.$$

Une suite de nombres vérifiant la condition de Cauchy est dite *suite de Cauchy* ou encore *suite fondamentale.*

La réciproque est vraie: *si une suite de nombres réels* $\{x_n\}$ *est une suite de Cauchy, alors elle est convergente, autrement dit il existe un nombre* $a$ *(fini) tel que* $x_n \to a$, $n \to \infty$.

DÉMONSTRATION. Commençons par prouver qu'une suite de Cauchy est bornée. Prenons $\varepsilon = 1$ et choisissons le nombre $n_0 = n_0(1)$ de la condition de Cauchy tel que

$$|x_n - x_m| < 1, \quad \forall n,\ m > n_0.$$

d'où

$$1 > |x_n - x_m| \geqslant |x_n| - |x_m|$$

ou

$$1 + |x_m| \geqslant |x_n|, \quad \forall n,\ m > n_0. \tag{1}$$

Figeons $m > n_0$ et posons

$$M = \max_{n \leqslant n_0} \{1 + |x_m|,\ |x_n|\}.$$

En vertu de (1) on a alors

$$M \geqslant |x_n|, \quad \forall n \in \mathrm{N},$$

ce qui prouve que la suite $\{x_n\}$ est bornée.

Le théorème de Bolzano-Weierstrass nous dit que de toute suite bornée $\{x_n\}$ on peut extraire une suite $\{x_{n_k}\}$ convergeant vers un nombre fini $a$, c'est-à-dire que

$$\lim_{k \to \infty} x_{n_k} = a.$$

Montrons que dans ce cas la suite $\{x_n\}$ converge aussi vers $a$.

En effet la condition de Cauchy qui est remplie par la suite $\{x_n\}$ dit que pour tout $\varepsilon > 0$ il existe un $n_0$ tel que

$$|x_n - x_m| < \varepsilon/2, \quad \forall\, n,\ m > n_0. \tag{2}$$

D'autre part, puisque $x_{n_k} \to a$, $k \to \infty$, on peut exhiber un $k_0$ tel que

$$|x_{n_k} - a| < \varepsilon/2, \quad \forall\, k > k_0.$$

Comme $n_k \to \infty$, pour $k \to \infty$, on peut trouver $k_1 > k_0$ tel que $n_{k_1} > n_0$. Donc

$$|x_{n_{k_1}} - a| < \varepsilon/2. \tag{3}$$

En vertu de (2) où l'on aura posé $m = n_{k_1}$, on déduit à partir de (3)

$$|x_n - a| = |x_n - x_{n_{k_1}} + x_{n_{k_1}} - a| \leqslant |x_n - x_{n_{k_1}}| + \\ + |x_{n_{k_1}} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \forall n > n_0,$$

ce qui prouve que la suite $\{x_n\}$ converge vers $a$.

Ainsi nous avons prouvé le

THÉORÈME (CRITÈRE DE CAUCHY). *Pour qu'une suite de nombres réels* $\{x_n\}$ *soit convergente, il est nécessaire et suffisant qu'elle soit une suite de Cauchy.*

## § 12. Complétude et continuité d'un ensemble de nombres réels

Dans les paragraphes précédents nous avons prouvé toute une série de propriétés des nombres réels dont les plus importantes sont :

1) La convergence de toute suite monotone bornée (théorème du § 5).

2) Le principe des segments emboîtés (théorème du § 7).

3) L'existence de la borne supérieure d'un ensemble borné (théorème du § 8).

4) La convergence d'une suite de Cauchy (critère de Cauchy, théorème du § 11).

La disparité de ces propriétés n'est qu'apparente, car en fait elles sont étroitement liées. On démontre sans peine en se servant des propriétés I à IV des nombres que les propositions 1) à 4) sont équivalentes, c'est-à-dire que l'une quelconque d'entre elles entraîne les trois autres. On a montré que de 1) (ou ce qui revient au même de la propriété V, cf. § 6, chap. 1) et des propriétés I à IV résultent 2), 3) et 4).

Les propriétés 1) à 4) s'appellent encore *propriétés de continuité* ou *de complétude* de l'ensemble **R** des réels.

Pour éclaircir leur rôle considérons l'ensemble **Q** des nombres rationnels. Les propriétés I à IV sont vérifiées pour les nombres rationnels. La propriété V et partant les propriétés 1) à 4) ne le sont pas en général.

Voyons ceci sur un exemple. Nous aurons besoin à cet effet de l'ensemble **R** des réels.

Considérons une forme décimale non périodique illimitée

$$a = a_0,\ a_1a_2a_3\ \ldots$$

Donc, $a$ est un nombre irrationnel, c'est-à-dire qu'il appartient à **R** mais pas à **Q**. La forme décimale de $a$ engendre une suite de troncatures

$$a^{(n)} = a_0,\ a_1a_2\ \ldots\ a_n \quad (n = 1,\ 2,\ \ldots)$$

(de nombres rationnels) croissante et majorée par le nombre entier $a_0 + 1$. La suite $\{a^{(n)}\}$ de nombres rationnels ne converge pas vers un nombre rationnel. En effet, on sait que $\{a^{(n)}\}$ converge vers $a$ (voir exemple 9 du § 1) qui est irrationnel et que cette limite est unique.

Nous avons montré que la propriété 1) n'est en général pas vérifiée dans **Q**.

On démontre sans peine qu'il en est de même des propriétés 2), 3) et 4).

L'ensemble des nombres réels est *complet*, car il est justiciable de la propriété 4) qui dit que toute suite de Cauchy converge vers un nombre réel.

L'ensemble **Q** des rationnels n'est pas complet, car les suites de Cauchy de nombres rationnels ne convergent pas toutes vers des nombres rationnels. En ajoutant à **Q** l'ensemble des nombres irrationnels, on obtient l'espace des réels qui, lui, est complet.

CHAPITRE 3

# FONCTION. LIMITE D'UNE FONCTION

## § 1. Fonction

Soit $E$ un ensemble de nombres. On dit que sur $E$ est définie une *fonction* si à tout élément $x \in E$ est associé un nombre $y \in \mathbf{R}$, et on note

$$y = f(x) \quad (x \in E). \tag{1}$$

Cette définition de la fonction a été proposée par N. Lobatchevski et Dirichlet [1]). L'ensemble $E$ s'appelle *ensemble de définition* de la fonction $f(x)$. On dit aussi qu'est donnée une *variable indépendante* $x$, appelée *argument*, qui parcourt l'ensemble $E$ et qu'à tout $x \in E$ est associée une valeur d'une autre variable $y$ appelée *fonction* ou *variable dépendante*.

On se sert aussi du langage géométrique pour exprimer la notion de fonction. Une fonction est définie par la donnée d'un ensemble $E$ de points $x$ de la droite numérique — le domaine de définition — et par une loi qui à tout point $x \in E$ associe un point $y = f(x)$.

Si l'on interprète une fonction comme une loi associant à chaque nombre $x \in E$ un nombre $y$, il suffit de la désigner par une seule lettre $f$. Le symbole $f(x)$ représente le nombre $y$ que la loi $f$ associe à $x \in E$. Si le nombre 1 appartient par exemple au domaine de définition $E$ de la fonction $f$, alors $f(1)$ est la *valeur de la fonction* $f$ au point $x = 1$. Si 1 n'appartient pas à $E$, on dit que la fonction $f$ *n'est pas définie* au point $x = 1$.

L'ensemble $E_1$ des valeurs $y = f(x)$, où $x \in E$, s'appelle *image* de l'ensemble $E$ par la fonction $f$. On écrit parfois $E_1 = f(E)$. Certaines précautions sont à prendre lors de l'usage de cette notation, afin d'éviter toute confusion avec la notation $y = f(x)$, où $x$ est un point de $E$ et $y$ son image par $f$. On dit encore que la fonction $f$ *applique* l'ensemble $E$ *sur* l'ensemble $E_1$.

Si l'image $E_1 = f(E) \subset A$, où $A$ est un ensemble de nombres généralement distinct de $E_1$, on dit alors que la fonction $f$ *applique* $E$ *dans* $A$.

---

[1]) Lobatchevski Nicolas, illustre mathématicien russe, fondateur de la géométrie non euclidienne (1792-1856). Dirichlet Gustav Lejeune, mathématicien allemand (1805-1859).

Etant données deux fonctions $f$ et $\varphi$ définies sur un même ensemble $E$, on appelle *somme* $f+\varphi$, *différence* $f-\varphi$, *produit* $f\varphi$ et *quotient* $f/\varphi$ les fonctions dont les valeurs sont exprimées respectivement par les formules

$$f(x)+\varphi(x), \quad f(x)-\varphi(x), \quad f(x)\,\varphi(x), \quad \frac{f(x)}{\varphi(x)} \quad (x \in E), \qquad (2)$$

dans le cas du quotient on admet que $\varphi(x) \neq 0$ sur $E$.

On se sert aussi des lettres $F$, $\Phi$, $\Psi$, . . . pour désigner des fonctions, et des lettres $z$, $u$, $v$, . . . pour désigner des variables.

Si une fonction $f$ applique un ensemble $E$ dans $E_1$ et une fonction $F$, l'ensemble $E_1$ dans un ensemble $E_2$, on dit que la fonction $z = F(f(x))$ est une *fonction de fonction*, ou une *fonction composée* de $f$ et de $F$. La composée est définie sur l'ensemble $E$ et applique $E$ dans $E_2$.

On définit par analogie la composée de $n$ fonctions: $z = F_1(F_2(F_3(\ldots(F_n(x))\ldots)))$.

Citons quelques exemples de fonctions. L'aire $S$ d'un disque est une fonction du rayon $r$ exprimée par la formule $S = \pi r^2$. Cette fonction est définie visiblement sur l'ensemble de tous les nombres $r > 0$.

On peut faire abstraction de l'aire du disque et considérer la dépendance de $S$ par rapport à $r$ exprimée par la formule $S = \pi r^2$. La fonction $S = \varphi(r)$ est définie sur l'axe des réels tout entier et pas seulement pour les $r > 0$.

Voici d'autres exemples de fonctions:

$$1)\ y=\sqrt{1-x^2}, \quad 2)\ x=\log(1+x), \quad 3)\ y=x'-1,$$

$$4)\ y=\frac{x^2-1}{x-1}, \quad 5)\ y=\operatorname{Arcsin} x.$$

Il s'agit de fonctions prenant des valeurs réelles $y$ pour des valeurs réelles de $x$. Il est immédiat de voir que ces fonctions sont définies respectivement sur:

1) l'intervalle fermé $[-1, 1]$;
2) l'intervalle $]-1, +\infty[$;
3) l'axe numérique tout entier;
4) l'axe numérique tout entier privé du point $x = 1$;
5) l'intervalle fermé $[-1, 1]$.

Les fonctions des exemples 1) et 2) peuvent être traitées comme des fonctions composées: 1) $y = \sqrt{u}$, $u = 1 - v$, $v = x^2$; 2) $y = \log u$, $u = 1 + x$.

Les graphes constituent un important procédé de définition des fonctions. Considérons un système de coordonnées rectangulaires $xOy$ (fig. 11), un intervalle fermé $[a, b]$ sur l'axe $Ox$, et traçons une courbe $\Gamma$ jouissant de la propriété suivante: toute droite parallèle à l'axe $Oy$ et passant par un point $x \in [a, b]$ coupe la courbe $\Gamma$

en un point $A$. La courbe $\Gamma$ s'appelle *graphe*. Ce graphe définit une fonction $y = f(x)$ sur l'intervalle $[a, b]$: si $x \in [a, b]$, la valeur $y = f(x)$ se définit comme l'ordonnée du point $A$ (fig. 11). Donc, le graphe établit bien une correspondance entre $x$ et $y = f(x)$.

Nous venons de définir une fonction par un graphe sur un ensemble $E$ qui est un intervalle fermé. Pour ensemble de définition $E$ on peut avoir un intervalle ouvert ou semi-ouvert, l'axe numérique tout entier, l'ensemble des points rationnels compris dans un intervalle donné, etc.

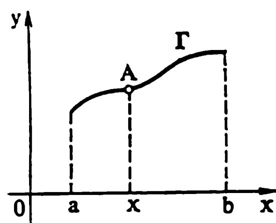



Fig. 11

Soient une fonction $f(x)$ définie sur un intervalle $]a, b[$ et un nombre $\alpha \neq 0$. On peut former plusieurs fonctions avec $\alpha$ et $f$: 1) $\alpha f(x)$; 2) $f(x) + \alpha$; 3) $f(x - \alpha)$; 4) $f(\alpha x)$. Les fonctions 1) et 2) sont définies sur le même intervalle $]a, b[$. Les graphes des fonctions 1) à 4) se déduisent de celui de $f(x)$ soit par une translation, soit par une homothétie.

On dit qu'une fonction est *paire* si $f(-x) = f(x)$ et *impaire* si $f(-x) = -f(x)$.

Le graphe d'une fonction paire est de toute évidence symétrique par rapport à l'axe $Oy$, celui d'une fonction impaire, symétrique par rapport à l'origine des coordonnées. Les fonctions $x^{2k}$ ($k$ est un entier naturel), $\cos x$, $\log |x|$, $\sqrt{1 + x^2}$, $f(|x|)$ sont paires, les fonctions $x^{2k+1}$ ($k \geqslant 0$ est un entier), $\sin x$, $x\sqrt{1 + x^2}$, $xf(|kx|)$, impaires.

Il est immédiat de voir que *le produit de deux fonctions paires ou de deux fonctions impaires est une fonction paire et que le produit d'une fonction paire par une fonction impaire est une fonction impaire.*

Il est entendu que la majorité des fonctions ne sont ni paires ni impaires.

On peut encore définir le graphe d'une fonction $y = f(x)$, $x \in E$, comme l'ensemble des points $(x, f(x))$.

On dit qu'une fonction $f$ est *strictement croissante* (resp. *croissante*) sur $E$ si pour tous $x_1, x_2 \in E$ tels que $x_1 < x_2$ on a $f(x_1) < f(x_2)$ (resp. $f(x_1) \leqslant f(x_2)$).

On dit qu'une fonction $f$ est *strictement décroissante* (resp. *décroissante*) sur $E$ si pour tous $x_1, x_2 \in E$ tels que $x_1 < x_2$, on a $f(x_1) > f(x_2)$ (resp. $f(x_1) \geqslant f(x_2)$).

On dit qu'une fonction $f$ est *bornée* (resp. *non bornée*) sur $E$ si l'image $E_1 = f(E)$ de $E$ par $f$ est un ensemble borné (resp. non borné).

Par exemple, la fonction $y = 1/x$ est décroissante et bornée sur l'intervalle $[1, \infty[$, mais pas sur l'intervalle $]0, \infty[$.

On dit qu'une fonction $f$ définie sur un ensemble $E$ est *périodiaue de période* $T > 0$ si $f(x) = f(x + T)$, $x, x + T \in E$.

Par exemple, la fonction $\sin x$ est de période $2\pi$, la fonction $\sin mx$, où $m \in \mathrm{N}$, de période $T = 2\pi/m$.

EXEMPLE 6. La fonction

$$y = \operatorname{sgn} x = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0 \end{cases}$$

est définie sur l'intervalle $]-\infty, \infty[$. Elle est impaire et prend trois valeurs: 1, 0, −1.

EXEMPLE 7. La fonction

$$y = \begin{cases} x^2 + 1, & x \leqslant 0, \\ \sin x, & x > 0 \end{cases}$$

est représentée par le graphe de la figure 12. Elle décroît sur $]-\infty, 0[$ et est $2\pi$-périodique sur l'intervalle $]0, \infty[$. Cette fonction est donnée par des formules différentes dans des régions différentes de son domaine de définition.

Une fonction peut être définie par un tableau. Si par exemple l'on mesure la température $T$ de l'air toutes les heures et que l'on consigne les résultats obtenus dans un tableau:

| $t$ | 0 | 1 | ... | 24 |
|---|---|---|---|---|
| $T$ | $T_0$ | $T_1$ | ... | $T_{24}$ |

on obtient une fonction $T = f(t)$ définie sur l'ensemble des entiers naturels compris entre 0 et 24.

Si une fonction $y = f(x)$ est définie sur un ensemble $E$ par une formule, il lui correspond un graphe qui la représente géométriquement. Mais si une fonction est définie par un graphe, peut-elle être exprimée par une formule? C'est un problème très épineux et avant de lui apporter une réponse il faut définir ce qu'on entend par formule. Quand nous avons dit qu'une fonction $y = f(x)$ est exprimée par une formule, nous avons tacitement admis que $y$ se déduit à partir de $x$ par un nombre fini d'opérations telles que l'addition, la soustraction, la multiplication, la division, l'extraction d'une racine carrée ou autre, la prise d'un logarithme, les opérations sin, cos, Arcsin et autres opérations algébriques et trigonométriques.

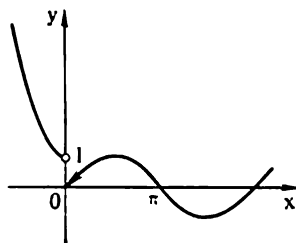


Fig. 12

L'analyse mathématique nous fournit des méthodes pour élargir

la notion de formule. L'une d'elles est le développement d'une fonction en une série de fonctions élémentaires.

La plupart ou voire même toutes les fonctions que l'on rencontre en pratique peuvent être représentées par une série illimitée dont les termes sont des fonctions élémentaires qui seront définies plus bas. Mais nous reviendrons sur ce problème.

Quoi qu'il en soit, qu'une fonction $f(x)$ soit définie par une formule ou par un autre procédé, par exemple par un graphe, elle peut être un objet d'étude par les méthodes de l'analyse mathématique si elle satisfait à certaines conditions subsidiaires, telles la continuité, la monotonie, la convexité, la dérivabilité, etc. Mais n'anticipons pas.

La notion de limite qui est fondamentale en analyse est un important outil d'étude d'une fonction. Ce chapitre lui sera consacré.

Si à un nombre $x$ d'un ensemble donné $E$ une loi associe un ensemble $e_x$ de nombres $y$, on dit qu'est définie une *correspondance* $y = f(x)$. Si pour chaque $x$ l'ensemble $e_x$ est composé au plus d'un élément, alors $y = f(x)$ est une fonction.

Comme exemples de correspondances citons $\sqrt{x}$, $\arcsin x$, $\operatorname{arctg} x$, . . .

La correspondance $\sqrt{x}$ est définie pour $x \geqslant 0$. Elle est bivalente pour $x > 0$, car à tout nombre $x > 0$ elle fait correspondre deux nombres opposés $+\sqrt{x}$ et $-\sqrt{x}$. Dans la suite, par $\sqrt[k]{x}$ ($k = 2, 3, \ldots$) on entendra, sauf mention expresse du contraire, la valeur arithmétique de la racine $k$-ième de $x$, c'est-à-dire un nombre positif dont la puissance $k$ est égale à $x$ (cf. § 8). La correspondance $\arcsin x$ est infinivalente. A toute valeur $x \in [-1, 1]$ elle associe une infinité de valeurs $y$ définies par la formule

$$y = (-1)^k \operatorname{Arcsin} x + k\pi \quad (k = 0, \pm 1, \pm 2, \ldots).$$

On définit de façon analogue des fonctions de deux, de trois et plus généralement de $n$ variables.

Définissons une fonction de deux variables. Soit un ensemble $E$ de couples $(x, y)$. Si à tout couple $(x, y) \in E$ une loi $f$ associe un nombre $z$, on dit qu'est définie sur l'ensemble $E$ une *fonction* $z = f(x, y)$ *de deux variables*.

Comme à chaque couple de nombres $(x, y)$ est associé un point d'abscisse $x$ et d'ordonnée $y$ dans un plan rapporté à un système de coordonnées cartésiennes et inversement à chaque point est associé un couple $(x, y)$, on peut dire que la fonction $f(x, y)$ est définie sur l'ensemble $E$ des points du plan.

Dans un espace à trois dimensions rapporté à un système de coordonnées rectangulaires $(x, y, z)$ le graphe de la fonction $z = f(x, y)$ est le lieu géométrique des points $(x, y, f(x, y))$ dont les projections $(x, y)$ appartiennent à l'ensemble de définition $E$ de $f$.

Par exemple, le lieu géométrique des points dont les coordonnées vérifient l'équation

$$z=\sqrt{1-x^2-y^2}, \quad x^2+y^2\leqslant 1,$$

est la demi-sphère supérieure de rayon 1 et de centre 0 (voir § 25, chap. 10).

On peut définir une fonction de trois variables dans le même esprit. Le domaine de définition sera un ensemble de triplets $(x, y, z)$, ou ce qui revient au même les points $(x, y, z)$ d'un espace à trois dimensions rapporté à un système de coordonnées cartésiennes.

Si à chaque triplet de nombres $(x, y, z) \in E$ une loi associe un nombre $u$, on dit alors que sur $E$ est définie une fonction $u = F(x, y, z)$.

De façon analogue, on peut considérer un ensemble $E$ de $n$-tuples $(x_1, \ldots, x_n)$, où $n$ est un entier naturel donné. Si à chaque $n$-tuple de $E$ une loi associe un nombre $z$, on dit que $z$ est une *fonction des variables* $x_1, \ldots, x_n$, définie sur l'ensemble $E$ et l'on note $z = F(x_1, \ldots, x_n)$.

On appelle *espace à n dimensions* l'ensemble de tous les $n$-tuples $(x_1, \ldots, x_n)$. Cet espace n'existe pas dans la réalité pour $n > 3$. Il n'empêche qu'il est aussi important que l'espace à trois dimensions.

Si deux fonctions $f$ et $\varphi$ de $n$ variables sont définies sur un même ensemble $E$ de points $(x_1, \ldots, x_n)$ d'un espace à $n$ dimensions, on peut définir la somme $f + \varphi$, la différence $f - \varphi$, le produit $f\varphi$ et le quotient $f/\varphi$ comme des fonctions définies sur $E$ à l'aide des égalités (2), où il faut remplacer $x$ par $(x_1, \ldots, x_n)$. On définit de façon naturelle les fonctions composées telles que $f(\varphi(x, y), \psi(x, y, z)) = F(x, y, z)$, où $(x, y, z)$ est un triplet de nombres appartenant à un certain ensemble.

On cite plus bas quelques exemples de fonctions de plusieurs variables définies par des formules élémentaires.

EXEMPLE 8. $u = Ax + By + Cz + D$, où $A$, $B$, $C$ et $D$ sont des constantes réelles, est une fonction linéaire de $(x, y, z)$. Elle est définie sur l'espace tout entier. La forme générale d'une fonction linéaire de $n$ variables $(x_1, \ldots, x_n)$ est $u = \sum_{i=1}^{n} a_i x_i + b$, où $a_1, \ldots, a_n, b$ sont des constantes données. Cette fonction est définie sur l'espace tout entier des points $(x_1, \ldots, x_n)$.

EXEMPLE 9. $z = \log \sqrt{1 - x^2 - y^2}$. C'est une fonction réelle définie sur un disque de rayon 1 et de centre (0, 0) privé de ses points frontières, c'est-à-dire du cercle de rayon 1 et de centre (0, 0). La fonction $z$ n'est pas définie pour ces points, car log 0 n'a pas de sens.

EXEMPLE 10. La fonction

$$z=f(x, y)=\begin{cases} 0 & \text{pour } y\geqslant 0, \\ 1 & \text{pour } y<0 \end{cases}$$

se représente graphiquement par deux demi-plans parallèles dont la disposition par rapport au système de coordonnées $(x, y, z)$ est évidente.

On peut définir une fonction d'une variable sous *forme implicite*

$$F(x, y) = 0, \tag{3}$$

où $F$ est une fonction de $x$ et de $y$.

Soit définie une fonction $F$ sur un ensemble $G$ de points $(x, y)$. La relation (3) définit un sous-ensemble $\Omega$ de $G$ sur lequel s'annule la fonction $F$. L'ensemble $\Omega$ peut éventuellement être vide. Supposons que $\Omega$ n'est pas vide et soit $E$ un ensemble (visiblement non vide) de valeurs $x$ auxquelles est associé au moins un $y$ tel que $(x, y) \in \Omega$. L'ensemble $E$ est donc l'ensemble de tous les points $x$ à chacun desquels est associé un ensemble non vide $e_x$ de points $y$ tels que $(x, y) \in \Omega$ ou ce qui revient au même tels que la relation (3) soit vérifiée pour le couple indiqué $(x, y)$. On définit ainsi une fonction $y = \varphi(x)$ sur l'ensemble $E$, qui en général est une correspondance. On dit alors que la fonction $\varphi$ est *implicitement définie* par (3). On a de toute évidence

$$F(x, \varphi(x)) \equiv 0, \quad \forall x \in E.$$

On peut définir par analogie une fonction $x = \psi(y)$ de $y$ implicitement par (3). On aura pour cette fonction :

$$F(\psi(y), y) \equiv 0, \qquad \forall y \in E_1,$$

où $E_1$ est un ensemble de nombres. On dit encore que la fonction $y = \varphi(x)$ (ou $x = \psi(y)$) est solution de l'équation (3). La fonction $x = \psi(y)$ s'appelle *fonction réciproque* de $y = \varphi(x)$.

EXEMPLE 11. La relation

$$x^2 + y^2 = r^2, \tag{4}$$

où $r > 0$, définit implicitement la correspondance

$$y = \pm\sqrt{r^2 - x^2} \quad (-r \leqslant x \leqslant r);$$

pour $x = \pm r$, c'est une fonction. On admet naturellement que cette correspondance se décompose en deux fonctions continues $y = +\sqrt{r^2 - x^2}$ et $y = -\sqrt{r^2 - x^2}$ $(-r \leqslant x \leqslant r)$. Leurs représentations graphiques sont deux demi-cercles d'un cercle de rayon $r$ et de centre $O$. Ce cercle est le lieu géométrique des points dont les coordonnées $(x, y)$ vérifient l'équation (4). En se servant de la formule (4) on peut former des fonctions (discontinues) vérifiant l'équation (4), par exemple

$$y = \begin{cases} +\sqrt{r^2 - x^2}, & -r \leqslant x < 0, \\ -\sqrt{r^2 - x^2}, & 0 \leqslant x \leqslant r. \end{cases}$$

## § 2. Limite d'une fonction

On dit qu'un nombre $A$ est la *limite d'une fonction $f$ en un point* $a$ si cette fonction est définie dans un voisinage de $a$, c'est-à-dire sur un intervalle $]c, d[$, où $c < a < d$, sauf éventuellement en $a$, et si pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ dépendant de $\varepsilon$ et tel que pour tous les $x$ vérifiant la relation $0 < |x - a| < \delta$, l'on ait

$$|f(x) - A| < \varepsilon.$$

On note ceci

$$\lim_{x \to a} f(x) = A \text{ ou } f(x) \to A \ (x \to a).$$

On peut aussi définir la limite d'une fonction en termes de limite d'une suite.

On dit qu'un nombre $A$ est la *limite d'une fonction $f$ en un point $a$* si cette fonction est définie dans un voisinage de $a$, sauf éventuellement en $a$, et si la suite $\{f(x_n)\}$ converge vers le point $A$, quelle que soit la suite $\{x_n\}$ convergeant vers $a$ et telle que $x_n \neq a$ pour tous les $n$. Donc

$$\lim_{\substack{x_n \to a \\ x_n \neq a}} f(x_n) = A.$$

Il est entendu que la variable $x_n$ tend vers $a$ en prenant des valeurs pour lesquelles $f(x)$ est définie.

Ces définitions sont équivalentes. Supposons en effet qu'une fonction ait une limite au sens de la première définition et soit donnée une variable $x_n$ tendant vers $a$ sans jamais prendre la valeur $a$. Considérons un nombre $\varepsilon$ et choisissons $\delta$ comme dans la première définition. Puis prenons un entier naturel $n_0$ tel que $|x_n - a| < \delta$ pour $n > n_0$. Alors

$$|f(x_n) - A| < \varepsilon \quad \text{pour } n > n_0,$$

ce qui exprime que la suite de nombres $\{f(x_n)\}$ converge vers $A$. Puisque cette propriété est vraie pour toute suite $\{x_n\}$ convergeant vers $a$, pourvu que $x_n \neq a$ et que tous les $x_n$ appartiennent au domaine de définition de $f$, on a démontré que la première définition entraîne la seconde.

Inversement, supposons que la fonction $f(x)$ a une limite au sens de la deuxième définition, mais pas au sens de la première. Cela signifie qu'il existe un $\varepsilon$ au moins, que nous désignons par $\varepsilon_0$, pour lequel on ne peut exhiber le $\delta$ nécessaire, c'est-à-dire que pour *tout* $\delta$ il existe parmi les $x$, tels que $0 < |x - a| < \delta$ au moins un, noté $x^{(\delta)}$ tel que $|f(x^{(\delta)}) - A| \geqslant \varepsilon_0$.

Pour $\delta$ nous prenons tous les nombres $\delta = 1/k$ ($k = 1, 2, \ldots$) et pour chacun d'eux trouvons un point $x_k = x^{(\delta)}$ tel que

$$0 < |x_k - a| < 1/k \quad (x_k \neq a)$$

et

$$|f(x_k) - A| \geqslant \varepsilon_0 \quad (k = 1, 2, \ldots).$$

De ces relations on voit que $x_k \to a$ ($x_k \neq a$), alors que $f(x_k)$ ne tend manifestement pas vers $A$. Donc, l'hypothèse que la seconde définition de la limite n'entraîne pas la première nous conduit à une contradiction.

Ceci achève la démonstration de l'équivalence des deux définitions.

L'expression *limite d'une fonction en un point a* est souvent remplacée par *limite d'une fonction lorsque x tend vers a* ou plus brièvement *limite d'une fonction pour* $x \to a$. Cette expression traduit mieux la notion de limite, car $\lim\limits_{x\to a} f(x)$ exprime le comportement de la fonction $f(x)$ dans un petit voisinage du point $a$ privé de $a$. Elle dit que si $x$ tend vers $a$ suivant une loi quelconque sans jamais égaler $a$, alors la valeur correspondante de $f(x)$ tend à son tour vers $A$, c'est-à-dire est aussi proche que l'on veut de $A$.

Exemple 1. Soit la fonction $f(x) = (x^2 - 4)/(x - 2)$. Elle est définie pour tous les $x \neq 2$. Essayons de trouver sa limite pour $x \to 2$. Pour tout $x \neq 2$, on a $\frac{x^2-4}{x-2} = x + 2$, et comme dans le calcul de la limite pour $x \to 2$ on ne tient nullement compte de la valeur de $f$ en $x = 2$, on obtient

$$\lim_{x\to 2} \frac{x^2-4}{x-2} = \lim_{x\to 2} (x+2).$$

Cette relation traduit le fait que si l'une des limites existe, l'autre existe aussi et lui est égale.

Pour trouver cette limite on dispose généralement les calculs de la manière suivante:

$$\lim_{x\to 2} \frac{x^2-4}{x-2} = \lim_{x\to 2} (x+2) = \lim_{x\to 2} x + 2 = 4.$$

Signalons que les fonctions $f(x) = (x^2 - 4)/(x - 2)$ et $\varphi(x) = x + 2$ sont différentes. La première n'est pas définie pour $x = 2$, tandis que la seconde l'est pour toutes les valeurs de $x$. Cependant, lorsqu'on calcule la limite de ces fonctions pour $x \to 2$, peu nous importe que ces fonctions soient définies ou non en $x = 2$, et comme $f(x) = \varphi(x)$ pour $x \neq 2$, il vient

$$\lim_{x\to 2} f(x) = \lim_{x\to 2} \varphi(x) = \varphi(2).$$

Exemple 2. Il est manifeste que $\lim\limits_{x\to 1} x^2 = 1$, car si $x_n \to 1$, $x_n \neq 1$, alors $\lim x_n^2 = \lim x_n \cdot \lim x_n = 1 \cdot 1 = 1$. On peut établir ce fait en utilisant le langage de ε et de δ. Considérons un intervalle contenant le point 1, par exemple $]1/2, 3/2[$. Pour tout $x \in ]1/2, 3/2[$

il est évident que

$$| x^2 - 1 | = | x + 1 | \, | x - 1 | \leqslant \frac{5}{2} \, | x - 1 |.$$

Prenons maintenant un $\varepsilon > 0$ quelconque et posons $\delta = \min\left\{\frac{1}{2}, \frac{2}{5}\varepsilon\right\}$. Alors pour tous les $x$ tels $| x - 1 | < \delta$, on aura

$$| x^2 - 1 | \leqslant \frac{5}{2} \cdot \frac{2}{5} \varepsilon = \varepsilon.$$

EXEMPLE 3. La fonction $\sin(1/x)$ est définie pour toutes les valeurs $x \neq 0$ et est impaire (son graphe est représenté sur la figure 13 pour $x > 0$). Elle est donc définie au voisinage du point 0 mais pas en 0. Cette fonction n'admet pas de limite pour $x \to 0$, car il existe une suite de valeurs non nulles $x_k = 2/\pi\,(2k + 1)$ $(k = 0, 1, 2, \ldots)$ tendant vers 0, pour laquelle

$$f(x_k) = (-1)^k$$

ne tend vers aucune limite pour $k \to \infty$.

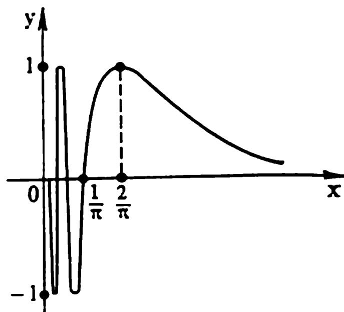



Fig. 13

Introduisons la définition suivante. On dit qu'un nombre $A$ est la *limite d'une fonction* $f(x)$ *pour* $x \to \infty$ et on note

$$A = \lim_{x \to \infty} f(x)$$

si $f$ est définie pour tous les $x$ tels que $| x | > K$ pour un certain $K > 0$ et si pour tout $\varepsilon > 0$ on peut exhiber un nombre $M > K$ tel que pour tous les $x$ vérifiant $| x | > M$ on ait $| f(x) - A | < \varepsilon$.

On démontre que cette définition est équivalente à la suivante.

Un nombre $A$ est la *limite d'une fonction* $f(x)$ *pour* $x \to \infty$ si la fonction $f(x)$ est définie pour tous les $x$ tels que $| x | > M$ pour un certain $M$ et

$$\lim_{x_n \to \infty} f(x_n) = A$$

pour toute suite $\{x_n\}$ convergeant vers $\infty$.

La démonstration de l'équivalence de ces deux définitions se fait de la même manière que pour la limite de $f$ en un point $a \neq \infty$.

En fait, de nombreuses propriétés des limites de $f(x)$ pour $x \to a$ ($a$ est un nombre fini) et pour $x \to \infty$ sont identiques. On peut exposer ces propriétés simultanément pour $x \to a$ et pour $x \to \infty$. On comprendra alors par la lettre $a$ soit un nombre fini, soit le symbole $\infty$. Si $a$ est un nombre, par voisinage du point $a$ on comprend tout intervalle $]c, d[$ contenant le point $a$. Donc, un *voisinage d'un point* $a$ est l'ensemble de tous les points $x \in\, ]c, d[$. Si $a = \infty$ ($+\infty$ ou $-\infty$), alors par voisinage de $a$ on convient d'entendre l'ensemble de tous

les $x$ tels que

$$|x| > M \quad (x > M, \text{ ou } x < -M, \ M > 0).$$

On écrira

$$\lim_{x \to a} f(x) = A,$$

où $a$ est un nombre fini ou infini, si la fonction $f(x)$ est définie dans un voisinage de $a$, sauf éventuellement en $a$ (cette restriction n'a de sens que si $a$ est fini), et si pour tout $\varepsilon > 0$ il existe un voisinage du point $a$ tel que pour tous les $x \neq a$ de ce dernier, l'on ait

$$|f(x) - A| < \varepsilon.$$

Cette définition combine visiblement les deux cas: le cas où $x$ tend vers un nombre fini $a$ et le cas où $x$ tend vers $+\infty$ ou $-\infty$.

Une fonction $f$ telle que $\lim_{x \to a} f(x) = 0$ s'appelle *infiniment petit pour* $x \to a$.

Passons à l'exposé des propriétés d'une fonction admettant une limite pour $x \to a$, où $a$ est un nombre fini ou infini. On désignera par $V(a)$ un voisinage arbitraire de $a$. Il est immédiat de vérifier que l'*intersection de deux voisinages* $V_1(a)$ *et* $V_2(a)$ *est encore un voisinage de* $a$.

THÉORÈME 1. *Si* $\lim_{x \to a} f(x) = A$, *où* $A$ *est un nombre fini, alors la fonction* $f(x)$ *est bornée dans un voisinage* $V(a)$, *c'est-à-dire qu'il existe un nombre strictement positif* $M$ *tel que*

$$|f(x)| \leqslant M \text{ pour tous les } x \in V(a), \quad x \neq a.$$

DÉMONSTRATION. Par hypothèse, il existe un voisinage $V(a)$ tel que

$$1 > |f(x) - A| \geqslant |f(x)| - |A| \quad (x \in V(a), \quad x \neq a).$$

D'où

$$|f(x)| \leqslant 1 + |A| = M,$$

c.q.f.d.

THÉORÈME 2. *Si* $\lim_{x \to a} f(x) = A$ *et* $A \neq 0$ *est un nombre fini, il existe alors un voisinage* $V(a)$ *tel que*

$$|f(x)| > |A|/2 \quad (x \in V(a), \quad x \neq a).$$

*De plus, pour les* $x$ *indiqués*

$$f(x) > A/2 \quad \text{si } A > 0,$$

*et*

$$f(x) < A/2 \quad \text{si } A < 0.$$

Démonstration. Par hypothèse, il existe pour un $\varepsilon = |A|/2$ un voisinage $V(a)$ tel que

$$|A|/2 > |A - f(x)| \geqslant |A| - |f(x)| \quad (x \in V(a), \quad x \neq a),$$

d'où $|f(x)| > |A|/2$ pour les $x$ indiqués. La première de ces inégalités équivaut aux suivantes:

$$A - |A|/2 < f(x) < A + |A|/2.$$

D'où pour $A > 0$

$$\frac{A}{2} = A - \frac{|A|}{2} < f(x),$$

et pour $A < 0$ .58

$$f(x) < A + \frac{|A|}{2} = \frac{A}{2},$$

c. q. f. d.

Théorème 3. *Si*

$$\lim_{x \to a} f_1(x) = A_1, \quad \lim_{x \to a} f_2(x) = A_2$$

*et*

$$f_1(x) \leqslant f_2(x)$$

*dans un voisinage* $V(a)$, $x \neq a$, *alors*

$$A_1 \leqslant A_2.$$

Démonstration. Supposons que $x_n \to a$, $x_n \neq a$; alors pour $n_0$ assez grand, on a

$$f_1(x_n) \leqslant f_2(x_n) \quad (n > n_0)$$

et après le passage à la limite, $A_1 \leqslant A_2$.

Théorème 4. *Si*

$$\lim_{x \to a} f_1(x) = A, \quad \lim_{x \to a} f_2(x) = A \tag{1}$$

*et*

$$f_1(x) \leqslant \varphi(x) \leqslant f_2(x) \tag{2}$$

*dans un voisinage* $V(a)$, $x \neq a$, *alors*

$$\lim_{x \to a} \varphi(x) = A. \tag{3}$$

Démonstration. Supposons que $x_n \to a$, $x_n \neq a$; alors pour $n_0$ assez grand, on a

$$f_1(x_n) \leqslant \varphi(x_n) \leqslant f_2(x_n) \quad (n > n_0)$$

et en vertu de (1) $\{\varphi(x_n)\}$ converge vers $A$. Comme la suite $\{x_n\}$ est une suite quelconque convergeant vers $a$, on a (3).

THÉORÈME 5 (CRITÈRE DE CAUCHY). *Pour qu'une fonction $f(x)$ admette une limite (finie) en un point $a$, il est nécessaire et suffisant qu'elle soit définie au voisinage de $a$, sauf éventuellement en $a$, et que pour tout $\varepsilon > 0$ il existe un voisinage $V(a)$ tel que*

$$|f(x') - f(x'')| < \varepsilon$$

*quels que soient $x'$, $x'' \in V(a)$, $x'$, $x'' \neq a$.*

DÉMONSTRATIPN. *Condition nécessaire.* Supposons que $\lim\limits_{x\to a} f(x) = A$, où $A$ est un nombre fini; il existe alors un voisinage de $a$ dans lequel $f(x)$ est définie sauf éventuellement en $a$. De plus pour tout $\varepsilon > 0$ on peut exhiber un voisinage $V(a)$ tel que pour $x \in V(a)$, $x \neq a$, l'on ait $|f(x) - A| < \varepsilon/2$. Soient $x'$, $x'' \in V(a)$, $x'$, $x'' \neq a$; alors

$$|f(x')-f(x'')| \leqslant |f(x')-A| + |A-f(x'')| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

ce qui prouve la condition nécessaire.

*Condition suffisante.* Supposons que la fonction $f(x)$ est définie au voisinage de $a$, sauf éventuellement en $a$ et que pour tout $\varepsilon > 0$ on peut exhiber un voisinage $V(a)$ de $a$ tel que $|f(x') - f(x'')| < \varepsilon$ pour tous les $x'$, $x'' \in V(a)$, $x'$, $x'' \neq a$. Considérons une suite quelconque $\{x_n\}$, $x_n \neq a$, convergeant vers $a$. Alors le critère de Cauchy pour les suites nous dit qu'il existe un nombre $n_0$ tel que pour $n$, $m > n_0$ l'on ait $x_n$, $x_m \in V(a)$. Dans ce cas

$$|f(x_n) - f(x_m)| < \varepsilon \qquad (n,\ m > n_0)$$

et la suite $\{f(x_n)\}$ vérifie le critère de Cauchy, donc est convergente.

Nous venons de prouver la propriété suivante: $\lim f(x_n)$ existe pour toute suite $\{x_n\}$ convergeant vers $a$ ($x_n \neq a$). De cette propriété il s'ensuit immédiatement que les limites $\lim f(x_n)$ correspondant à diverses suites convergeant vers $a$ sont égales. En effet, soient $x_n \to a$, $x'_n \to a$; $x_n$, $x'_n \neq a$ ($n = 1, 2, \ldots$). D'après ce que nous venons de démontrer il existe des nombres $A$ et $A'$ tels que $f(x_n) \to A$ et $f(x'_n) \to A'$. Formons une nouvelle suite: $\{x_1, x'_1, x_2, x'_2, x_3, \ldots\}$. Cette suite converge vers $a$. D'après ce qui a été démontré plus haut, la suite $\{f(x_1), f(x'_1), f(x_2), f(x'_2), \ldots\}$ doit aussi converger vers un certain nombre. Or, ceci n'a lieu que si $A = A'$, c.q.f.d.

THÉORÈME 6. *Si*

$$\lim_{x\to a} f(x) = A, \quad \lim_{x\to a} \varphi(x) = B,$$

*où $A$ et $B$ sont des nombres finis, alors*

$$\lim_{x\to a} [f(x) \pm \varphi(x)] = A \pm B, \quad \lim_{x\to a} [f(x)\, \varphi(x)] = AB$$

*et*

$$\lim_{x\to a} \frac{f(x)}{\varphi(x)} = \frac{A}{B}$$

*sous réserve que $B \neq 0$.*

Prouvons la deuxième égalité à titre d'exemple. Soit $x_n \to a$, $x_n \neq a$ $(n = 1, 2, \ldots)$; alors

$$\lim f(x_n) = A, \quad \lim \varphi(x_n) = B.$$

Comme la limite d'un produit est égal au produit des limites, on obtient

$$\lim [f(x_n)\,\varphi(x_n)] = \lim f(x_n) \lim \varphi(x_n) = AB.$$

Cette égalité est valable pour toute suite $x_n \to a$, $x_n \neq a$, donc $\lim\limits_{x\to a} [f(x)\,\varphi(x)] = AB$.

Par définition, $\lim\limits_{x\to a} f(x) = \infty$ si la fonction $f(x)$ est définie dans un voisinage de $a$, sauf éventuellement en $a$, et si pour tout nombre strictement positif $M$ il existe un voisinage $V(a)$ de $a$ tel que

$$|f(x)| > M \quad (x \in V(a), \quad x \neq a).$$

On appelle *infiniment grand pour $x \to a$* une fonction $f(x)$ telle que $\lim\limits_{x\to a} f(x) = \infty$.

Si $\lim\limits_{x\to a} f(x) = \infty$ et si dans un voisinage de $a$ $f(x) > 0$ (resp. $f(x) < 0$), on écrit encore $\lim\limits_{x\to a} f(x) = +\infty$ (resp. $\lim\limits_{x\to a} f(x) = -\infty$).

Les théorèmes suivants sont immédiats.

THÉORÈME 7. *Si $f(x)$ et $\varphi(x)$ sont deux fonctions telles que dans un voisinage d'un point $a$*

$$|f(x)| > M > 0$$

*et*

$$\lim_{x\to a} \varphi(x) = 0 \quad (\varphi(x) \neq 0 \text{ pour } x \neq a),$$

*alors*

$$\lim_{x\to a} \frac{f(x)}{\varphi(x)} = \infty.$$

THÉORÈME 8. *Si $\lim\limits_{x\to a} f(x) = A$ et $\lim\limits_{x\to a} \varphi(x) = \infty$, alors*

$$\lim_{x\to a} \frac{f(x)}{\varphi(x)} = 0.$$

COROLLAIRE. *Si $\varphi(x) \to 0$, $x \to a$ $(\varphi(x) \neq 0)$, alors*

$$\lim_{x\to a} \frac{1}{\varphi(x)} = \infty,$$

*et si* $\varphi(x) \to \infty$, $x \to a$ $(\varphi(x) \neq 0)$, *alors*

$$\lim_{x \to a} \frac{1}{\varphi(x)} = 0.$$

On peut encore définir la limite à droite et à gauche d'une fonction $f$ en un point $a$.

On dit qu'un nombre $A$ est la *limite à droite* (resp. *à gauche*) *d'une fonction $f$ en un point $a$* si cette fonction est définie sur un intervalle semi-ouvert $]a, b]$ (resp. $[b, a[$) et si

$$\lim_{\substack{x_n \to a \\ x_n > a}} f(x_n) = A \quad (\text{resp. } \lim_{\substack{x_n \to a \\ x_n < a}} f(x_n) = A)$$

pour toute suite $\{x_n\}$.

On désigne la limite à droite (resp. à gauche) d'une fonction $f$ en $a$ par:

$$f(a+0) = \lim_{\substack{x \to a \\ x > a}} f(x), \tag{4}$$

$$f(a-0) = \lim_{\substack{x \to a \\ x < a}} f(x). \tag{5}$$

Si $f$ est définie sur l'intervalle ouvert $]a, b[$, alors $f(a + 0)$ n'a de sens qu'en $a$ et $f(b - 0)$, qu'en $b$.

Remarque. La relation

$$f(a + 0) = f(a - 0) = A \tag{6}$$

équivaut à l'existence de la limite

$$\lim_{x \to a} f(x) = A. \tag{7}$$

En effet, on peut exprimer (6) comme suit: $\forall \varepsilon > 0$, $\exists \delta > 0$: $|f(x) - A| < \varepsilon$, $\forall x$: $0 < |x - a| < \delta$, $x > a$; $|f(x) - A| < \varepsilon$, $\forall x$: $0 < |x - a| < \delta$, $x < a$, ou de façon plus concise: $\forall \varepsilon > 0$, $\exists \delta > 0$: $|f(x) - A| < \varepsilon$, $\forall x: 0 < |x - a| < \delta$, ce qui est équivalent à (7).

## § 3. Continuité d'une fonction

La figure 14 représente le graphe d'une fonction $y = f(x)$ $(x \in [a, b])$. Soit un point $x \in [a, b]$. Tout point $x' \in [a, b]$ proche de $x$ peut être représenté par la somme $x' = x + \Delta x$, où $\Delta x$ est un nombre positif ou négatif appelé *accroissement* de $x$. La différence

$$\Delta f = \Delta y = f(x + \Delta x) - f(x)$$

s'appelle *accroissement de la fonction* $f$ en $x$, correspondant à l'accroissement $\Delta x$. Observons que $\Delta x$ est tel que $x + \Delta x \in [a, b]$. Sur la figure 14 l'accroissement $\Delta y$ est égal à $BC$.

Faisons tendre $\Delta x$ vers 0; il est évident que

$$\Delta y \to 0 \quad (\Delta x \to 0). \tag{1}$$

Considérons maintenant la courbe de la figure 15. Elle présente un saut en $x_0$. On dit alors que cette courbe est discontinue. Pour qu'un graphe représente une fonction $y = F(x)$ en $x_0$, on conviendra que $F(x_0)$ est égal à la longueur du segment $Ax_0$. Pour spécifier

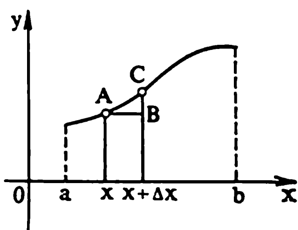



Fig. 14

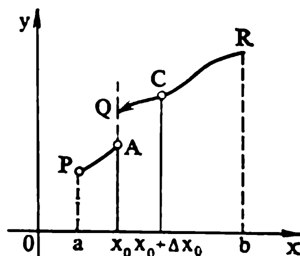



Fig. 15

ceci, le point $A$ est représenté par un rond sur le graphe, alors que la flèche indique que le point $Q$ n'appartient pas à ce graphe.

Donnons à $x_0$ un accroissement $\Delta x_0$ et définissons l'accroissement correspondant de la fonction:

$$\Delta F = F(x_0 + \Delta x_0) - F(x_0).$$

Si $\Delta x_0 \to 0$, on ne peut pas affirmer que $\Delta F \to 0$: $\Delta F$ tend vers 0 ou vers un nombre égal à la longueur du segment $AQ$ selon que $\Delta x_0$ tend vers 0 par valeurs négatives ou positives.

On dit qu'*une fonction* $f$ *définie sur un intervalle* $[a, b]$ *est continue en un point* $x \in [a, b]$ *si* $\Delta f \to 0$ *lorsque* $\Delta x \to 0$ *quel que soit le mode de tendance de* $\Delta x$ *vers* 0. La propriété de continuité est traduite par l'expression (1) ou encore par la relation:

$$\lim_{\Delta x \to 0} \Delta y = 0, \tag{2}$$

qui se lit « la limite de $\Delta y$ est nulle lorsque $\Delta x$ tend vers 0 suivant une loi quelconque ». Cependant on omet généralement l'expression « suivant une loi quelconque ».

Si une fonction $f$ définie sur $[a, b]$ n'est pas continue en un point $x \in [a, b]$, c'est-à-dire si elle n'est pas justiciable de la relation (2) pour au moins une méthode de tendance de $\Delta x$ vers 0, on dit alors qu'elle est *discontinue* en $x$.

La fonction représentée sur la figure 14 est continue en tout point $x \in [a, b]$, celle de la figure 15, l'est visiblement en tout point $x \in [a, b]$ sauf en $x_0$, car en ce point la relation (2) est mise en défaut lorsque $\Delta x$ tend vers 0 par valeurs positives.

Une fonction continue en tout point d'un intervalle est dite *continue sur* cet *intervalle.*

La continuité d'une fonction est une propriété à laquelle nous aurons souvent affaire en pratique, une propriété qui traduit le fait qu'à tout petit accroissement de la variable indépendante correspond un petit accroissement de la fonction. Un remarquable exemple de fonction continue nous est donné par les diverses lois de mouvement des corps $s = f(t)$ qui expriment le chemin $s$ parcouru en fonction du temps $t$. Le temps et l'espace sont continus et les diverses lois de mouvement $s = f(t)$ établissent entre $s$ et $t$ une relation continue caractérisée par le fait qu'à tout petit accroissement du temps correspond un petit accroissement du chemin.

L'observation des milieux continus — solides, liquides ou gazeux — a conduit l'homme à faire abstraction de la continuité. En effet, tout milieu physique est un amas de particules en mouvement. Ces particules et les distances qui les séparent sont si petites en regard des dimensions des milieux auxquels on a affaire en étudiant les phénomènes physiques macroscopiques que l'on peut considérer approximativement que la masse du milieu envisagé est continûment distribuée dans l'espace occupé. De nombreuses disciplines, telles la dynamique des fluides, l'aérodynamique et la théorie de l'élasticité se servent de cette hypothèse. La notion mathématique de continuité joue naturellement un grand rôle dans ces disciplines, de même d'ailleurs que dans beaucoup d'autres.

Les fonctions continues constituent la plus importante classe de fonctions manipulées par l'analyse mathématique.

Comme exemples de fonctions continues citons les fonctions élémentaires (voir § 8) qui sont continues sur leurs intervalles de définition.

Les fonctions discontinues décrivent des processus naturels discontinus. La vitesse d'un corps varie discontinûment sous l'effet d'un choc. De nombreuses transitions sont accompagnées de sauts. Ainsi, la relation $Q = f(t)$ entre la température $t$ d'un gramme d'eau (de glace) et le nombre de calories $Q$, lorsque $t$ varie entre $-10°$ et $+30°$, si l'on convient que $Q = 0$ pour $t = -10°$, s'exprime par

$$Q(t) = \begin{cases} 0{,}5t + 5, & -10 \leqslant t < 0, \\ t + 85, & 0 < t \leqslant 30. \end{cases}$$

On admet que la capacité calorifique de la glace est égale à 0,5. Pour $t = 0$, cette fonction est indéfinie (elle prend des valeurs distinctes) ; on peut par souci de commodité convenir que pour $t = 0$ elle prend

une valeur bien définie, par exemple $f(0) = 45$. La fonction $Q = f(t)$ est visiblement discontinue pour $t = 0$ (fig. 16).

Définissons la continuité d'une fonction $f$ en un point.

*On dit qu'une fonction $f(x)$ est continue en un point $x_0$ si elle est définie dans un voisinage de $x_0$ de même qu'en $x_0$ et si son accroissement en ce point tend vers* 0 *avec $\Delta x$:*

$$\lim_{\Delta x\to 0} \Delta y = \lim_{\Delta x\to 0} [f(x_0 + \Delta x) - f(x_0)] = 0. \tag{3}$$

Si l'on pose $x = x_0 + \Delta x$, on obtient la définition équivalente suivante: *on dit qu'une fonction $f$ est continue en un point $x_0$ si elle*

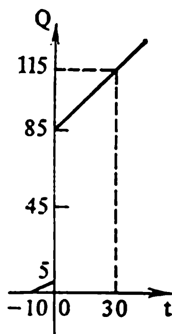



Fig. 16

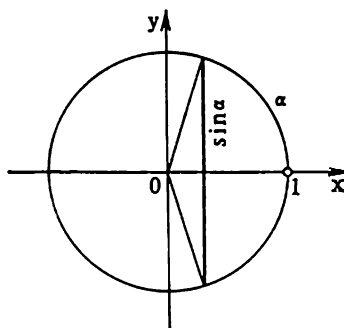



Fig. 17

*est définie dans un voisinage de $x_0$, y compris en $x_0$, et si*

$$\lim_{x\to x_0} f(x) = f(x_0); \tag{4}$$

ou dans le langage de $\varepsilon$ et $\delta$: *si pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que*

$$|f(x) - f(x_0)| < \varepsilon, \quad \forall x: |x - x_0| < \delta.$$

La relation (4) peut encore s'écrire:

$$\lim_{x\to x_0} f(x) = f\left(\lim_{x\to x_0} x\right). \tag{4'}$$

Cette relation montre qu'on *peut passer à la limite sous le signe d'une fonction continue.*

Exemple 1. La constante $y = C$ est une fonction continue en tout point $x$. En effet, en $x$ et $x + \Delta x$ on a $y = C$. Donc $\Delta y = y(x + \Delta x) - y(x) = C - C = 0$ et

$$\lim_{\Delta x\to 0} \Delta y = \lim_{\Delta x\to 0} 0 = 0.$$

EXEMPLE 2. La fonction $y = x$ est continue pour tout $x$, car $\Delta y = \Delta x$ et par suite $\Delta y \to 0$ avec $\Delta x$.

EXEMPLE 3. La fonction $y = \sin x$ est continue pour tout $x$. En effet

$$|\Delta y| = |\sin(x+\Delta x) - \sin x| = \left| 2 \sin \frac{\Delta x}{2} \cos\left(x + \frac{\Delta x}{2}\right)\right| \leqslant \leqslant 2\,|\sin(\Delta x/2)|. \quad (5)$$

Or, pour tout $\alpha$ on a

$$|\sin \alpha| < |\alpha|. \quad (6)$$

Si $0 < \alpha < \pi/2$, ceci résulte de la figure 17 (l'arc de longueur $2\alpha$ est plus grand que la corde de longueur $2\sin\alpha$ qui le sous-tend). Pour $\alpha = 0$, l'inégalité (6) se transforme en égalité. Si $0 < |\alpha| < \pi/2$, alors $|\sin \alpha| = \sin|\alpha| \leqslant |\alpha|$. Enfin, si $|\alpha| > \pi/2$, alors $|\sin \alpha| \leqslant 1 < \pi/2 \leqslant |\alpha|$. De (5) et de (6) on déduit que

$$|\Delta y| \leqslant 2\left|\sin \frac{\Delta x}{2}\right| \leqslant 2\,\frac{|\Delta x|}{2} = |\Delta x|,$$

c'est-à-dire que

$$|\Delta y| \leqslant |\Delta x|.$$

Il est alors évident que

$$\lim_{\Delta x \to 0} \Delta y = 0.$$

On peut encore dire que pour tout $\varepsilon > 0$ il existe un $\delta > 0$, en l'occurrence $\delta = \varepsilon$ tel que

$$|\Delta y| < \varepsilon, \quad \forall \Delta x: \quad |\Delta x| < \delta = \varepsilon.$$

Signalons un théorème important.

THÉORÈME 1. *Si deux fonctions $f$ et $\varphi$ sont continues en un point $x = a$, il en est de même de leur somme, de leur différence, de leur produit et de leur quotient* ($\varphi(a) \neq 0$).

Ce théorème est une conséquence immédiate du théorème 6, § 2, puisque

$$f(a) = \lim_{x \to a} f(x), \quad \varphi(a) = \lim_{x \to a} \varphi(x).$$

Signalons un autre théorème important de continuité d'une fonction composée.

THÉORÈME 2. *Soient données une fonction $f(u)$ continue en un point $u = A$ et une fonction $u = \varphi(x)$ continue en un point $x = a$, et admettons que $\varphi(a) = A$. Alors la fonction composée $F(x) = f[\varphi(x)]$ est continue au point $x = a$.*

Démonstration. On remarquera que de la définition de la continuité de la fonction $f$ en $A$ il s'ensuit qu'elle est définie dans un voisinage de $A$. Donc

$$\lim_{x\to a} F(x) = \lim_{x\to a} f[\varphi(x)] = \lim_{u\to A} f[u] = f(A) = f[\varphi(a)] = F(a).$$

On a effectué la substitution $u = \varphi(x)$ et utilisé la continuité de $\varphi$ en $x = a$: $\varphi(x) \underset{x\to a}{\to} \varphi(a) = A$.

Exemple 4. La fonction

$$P(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n,$$

où $a_k$ sont des coefficients constants, s'appelle *polynôme* de degré $n$. Cette fonction est continue pour tout $x$. En effet, $P(x)$ s'obtient à partir des constantes $a_0, \ldots, a_n$ et de la fonction $x$ par un nombre fini d'opérations arithmétiques: l'addition, la soustraction et la multiplication. Or, une constante est une fonction continue (voir exemple 1) et la fonction $x$ est aussi continue (voir exemple 2), donc en vertu du théorème 1 la fonction $P(x)$ est aussi continue.

Exemple 5. La fonction $y = \cos x$ est continue, car elle est la composée de deux fonctions continues: $y = \sin u$, $u = \pi/2 - x$.

Exemple 6. La fonction

$$y = \operatorname{tg} x = \frac{\sin x}{\cos x}, \quad x \neq \frac{\pi}{2} + k\pi \quad (k = 0, \pm 1, \pm 2, \ldots),$$

est continue pour les $x$ indiqués, car (cf. théorème 1) elle est égale au quotient d'une fonction continue par une fonction continue non nulle.

Exemple 7. La fonction

$$y = \sin^3 x^5$$

est continue pour tout $x$, car elle est la composée de fonctions continues: $y = u^3$, $u = \sin v$, $v = x^5$ (cf. théorème 2).

Exemple 8. La fonction $y = |x|$ est continue pour tout $x$, car

$$|\Delta y| = ||x + \Delta x| - |x|| \leqslant |x + \Delta x - x| = |\Delta x| \to 0$$
$$\text{pour } \Delta x \to 0.$$

Exemple 9. Si une fonction $f(x)$ est continue en un point $x_0$, il en est de même de la fonction $|f(x)|$.

Ceci résulte du théorème 2 et de l'exemple 8, puisque la fonction $|f(x)|$ est la composée de deux fonctions continues: $y = |u|$, $u = f(x)$.

Signalons encore deux théorèmes qui résultent immédiatement des théorèmes 1 et 2 du § 2, relatifs à la limite de fonctions.

Théorème 3. *Si une fonction $f$ est continue en un point $a$, il existe un voisinage $V(a)$ dans lequel elle est bornée.*

THÉORÈME 4. *Si une fonction f est continue en un point a et* $f(a) \neq 0$, *il existe un voisinage V (a) dans lequel*

$$|f(x)| > |f(a)|/2.$$

*Si en outre* $f(a) > 0$, *alors*

$$f(a)/2 < f(x) \quad (x \in V(a)),$$

*et si* $f(a) < 0$, *alors*

$$f(x) < f(a)/2 \quad (x \in V(a)).$$

## § 4. Discontinuités de première et de seconde espèces

Par définition *une fonction f est continue en un point* $x = a$ *à droite* (resp. *à gauche*) *si*

$$f(a) = f(a + 0) \quad (\text{resp. } f(a) = f(a - 0))$$

(cf. fin du § 2).

On peut encore définir la continuité d'une fonction $f$ en un point $a$ de la manière suivante: *on dit qu'une fonction f est continue en un*

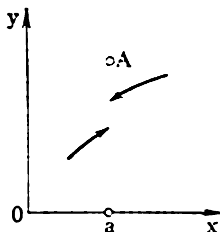


Fig. 18

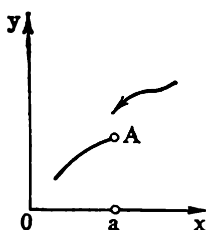


Fig. 19

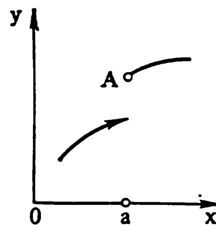


Fig. 20

*point* $x = a$ *si elle est définie dans un voisinage de a, y compris en a, et si existent des limites* $f(a + 0)$ *et* $f(a - 0)$ *telles que*

$$f(a) = f(a + 0) = f(a - 0). \tag{1}$$

Si une fonction $f$ est telle qu'elle admet les limites $f(a + 0)$ et $f(a - 0)$, et $f(a + 0) \neq f(a - 0)$, alors elle est *discontinue* en $a$, et cette discontinuité est dite *discontinuité de première espèce.*

Les figures 18 à 23 représentent six graphes de fonctions présentant une discontinuité de première espèce en $a$. La lettre $A$ désigne le point $A = (a, f(a))$. La pointe de la flèche n'appartient pas au graphe.

Les figures 18 à 21 représentent les graphes de fonctions pour lesquelles les nombres $f(a)$, $f(a + 0)$ et $f(a - 0)$ ont un sens. Sur

le graphe de la figure 18 les nombres $f(a)$, $f(a+0)$ et $f(a-0)$ sont deux à deux distincts : la fonction est discontinue en $a$ à gauche et à droite. La figure 19 représente le graphe d'une fonction continue en $a$ à gauche et discontinue à droite. Sur la figure 21 on a $f(a+0) = f(a-0) \neq f(a)$. On dit alors que la fonction $f$ présente en $a$ *une discontinuité artificielle*: en effet on peut éliminer cette discontinuité en $a$ en posant $f(a) = f(a+0) = f(a-0)$. La fonction

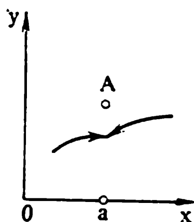



Fig. 21

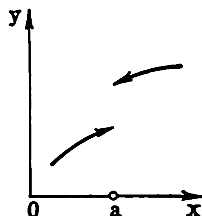



Fig. 22

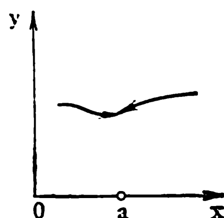



Fig. 23

représentée sur la figure 22 n'est pas définie en $a$. La fonction représentée sur la figure 23 n'est pas définie non plus en $a$, mais $f(a+0) = f(a-0)$, donc si l'on complète cette fonction en $a$ en posant $f(a) = f(a+0) = f(a-0)$, on rend la fonction $f$ continue en $a$.

Dans les figures 22 et 23 la fonction $f$ est définie au voisinage de $a$ mais pas en $a$. On dit alors que la fonction $f$ est discontinue en $a$, bien que l'idée de continuité et de discontinuité en $a$ consiste à comparer $f(a)$ et $f(x)$ pour les valeurs de $x$ proches de $a$.

On dit qu'une fonction $f$ présente une *discontinuité de seconde espèce* en un point $a$ si elle ne possède pas de limite à droite ou à gauche en $a$, ou bien si elle ne possède de limite ni à droite ni à gauche en $a$, ou encore si elle possède des limites infinies à droite et à gauche.

Exemple 1. La fonction

$$f(x) = \begin{cases} \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

ne possède pas de limite à droite et à gauche au point $x = 0$ (voir exemple 3, § 2,). Donc, elle présente une discontinuité de seconde espèce au point $x = 0$.

Exemple 2. La fonction

$$\operatorname{sgn} x = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0, \end{cases}$$

est continue pour $x \neq 0$ et présente une discontinuité de première espèce au point $x = 0$. De plus, $\operatorname{sgn}(0 + 0) = 1$, $\operatorname{sgn}(0 - 0) = = -1$.

EXEMPLE 3. La fonction $[x]$, ou partie entière de $x$, est représentée sur la figure 24 pour $x \geqslant 0$. Elle est continue pour les $x$ non entiers. Si $x$ est entier, alors $[x + 0] = = x = [x]$ et $[x - 0] = x - 1$, donc elle présente une discontinuité de première espèce.

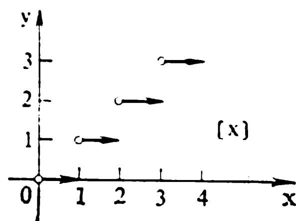



Fig. 24

EXEMPLE 4. La fonction

$$y = \begin{cases} 1/x, & x \neq 0, \\ 2, & x = 0, \end{cases}$$

est continue pour $x \neq 0$. Les limites à droite et à gauche en $x = 0$ sont infinies, donc cette fonction présente une discontinuité de seconde espèce en $x = 0$. On dit encore que cette fonction présente un *saut infini en* 0.

THÉORÈME 1. *Si une fonction $f$ est croissante sur un intervalle $[a, b]$, elle possède les limites $f(a + 0) \geqslant f(a)$ et $f(b - 0) \leqslant f(b)$.*

DÉMONSTRATION. Par hypothèse,

$$f(x) \leqslant f(b), \quad \forall x \in [a, b[,$$

autrement dit, $f$ est majorée par $f(b)$ sur $[a, b[$. Donc, $f$ admet une borne supérieure sur $[a, b[$ :

$$\sup_{x \in [a, b[} f(x) = M \leqslant f(b).$$

La propriété de la borne supérieure dit que pour tout $\varepsilon > 0$ on peut exhiber un $x_0 \in [a, b[$ tel que

$$M - \varepsilon < f(x_0) \leqslant M, \tag{2}$$

et puisque $f$ est une fonction croissante,

$$f(x_0) \leqslant f(x), \quad \forall x : x_0 < x < b. \tag{3}$$

De (2) et (3) il résulte que

$$M - \varepsilon < f(x) \leqslant M, \quad \forall x : x_0 < x < b,$$

ce qui prouve que la fonction $f$ admet une limite à gauche en $b$:

$$\lim_{\substack{x \to b \\ x < b}} f(x) = f(b - 0) = M \leqslant f(b).$$

De façon analogue, en considérant l'inégalité $f(a) \leqslant f(x)$ pour $x \in ]a, b]$, on prouve l'existence de

$$f(a+0) = \inf_{x \in ]a, b]} f(x) \geqslant f(a).$$

COROLLAIRE. *Si une fonction $f$ est croissante sur un intervalle $[a, b]$, alors en tout point $x \in [a, b[$ elle possède la limite à droite $f(x+0) \geqslant \geqslant f(x)$ et en tout point $x \in ]a, b]$, la limite à gauche $f(x-0) \leqslant \leqslant f(x)$.*

En effet, cette proposition est prouvée dans le théorème 1 pour $x = a$ et $x = b$. Supposons que $x \in ]a, b[$. La fonction $f$ est croissante sur les intervalles $[a, x]$ et $[x, b]$, donc d'après le théorème 1 elle possède les limites $f(x-0)$, $f(x+0)$, et $f(x-0) \leqslant f(x) \leqslant \leqslant f(x+0)$.

Il est évident que, pour que la fonction $f$ soit continue en $x$, il est nécessaire et suffisant que $f(x-0) = f(x+0)$.

Si $f(x-0) < f(x+0)$, alors la fonction $f$ présente une discontinuité de première espèce.

THÉORÈME 2. *L'ensemble des points de discontinuité d'une fonction $f$ monotone sur un intervalle $[a, b]$ est au plus dénombrable.*

DÉMONSTRATION. Supposons que la fonction $f$ possède plus d'un point de discontinuité et soient $x'$ et $x''$ ($x' < x''$) deux d'entre eux. Comme

$$f(x'+0) = \inf_{y \in ]x', x''[} f(y), \quad f(x''-0) = \sup_{y \in ]x', x''[} f(y),$$

alors

$$f(x'+0) \leqslant f(x''-0)$$

et les intervalles $]f(x'-0), f(x'+0)[$, $]f(x''-0), f(x''+0)[$ de l'axe $Oy$ sont disjoints.

A chaque point de discontinuité $x'$ de la fonction $f$ correspond un intervalle $]f(x'-0), f(x'+0)[$. Choisissons un point rationnel $\alpha_{x'}$ dans cet intervalle. D'après ce qui vient d'être dit il est clair qu'à divers points de discontinuité $x'$ correspondent divers points $\alpha_{x'}$. Or, l'ensemble des points rationnels est dénombrable. Donc, l'ensemble de tous les points $\alpha_{x'}$, de même que l'ensemble des points $x'$, est au plus dénombrable. C.q.f.d.

## § 5. Fonctions continues sur un intervalle fermé

On dit qu'une fonction $f$ est *continue sur un intervalle fermé* $[a, b]$ si elle l'est en tous les points de l'intervalle $]a, b[$, à droite au point $a$ et à gauche au point $b$.

Les fonctions continues sur des intervalles fermés sont douées de propriétés remarquables que nous nous proposons d'exposer.

Commençons par formuler des théorèmes traduisant ces propriétés, puis illustrons-les sur des graphes et des exemples avant de les prouver formellement.

THÉORÈME 1. *Si une fonction $f$ est continue sur un intervalle fermé $[a, b]$, alors elle est bornée sur $[a, b]$, c'est-à-dire qu'il existe une constante $K > 0$ telle que*

$$|f(x)| \leqslant K, \quad \forall x \in [a, b].$$

La figure 25 représente le graphe $\Gamma$ d'une fonction $f$ continue sur l'intervalle $[a, b]$. Il est évident qu'il existe un nombre $K > 0$ tel que $\Gamma$ est situé au-dessous de la droite $y = K$, mais au-dessus de la droite $y = -K$. C'est ce qu'exprime le théorème 1.

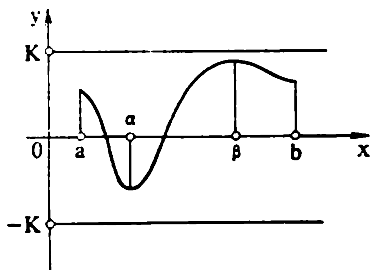


Fig. 25

On remarquera que si une fonction est continue sur un intervalle ouvert $]a, b[$ ou sur un intervalle semi-ouvert $[a, b[$ ou $]a, b]$, elle n'est pas forcément bornée. Exemple : la fonction $1/x$ est continue mais pas bornée sur l'intervalle $]0, 1]$.

Si l'on complète cette fonction en posant $f(0) = 0$, elle sera finie en tout point de l'intervalle $[0, 1]$, mais pas bornée sur cet intervalle.

THÉORÈME 2 (WEIERSTRASS). *Si une fonction $f$ est continue sur $[a, b]$, alors elle atteint son minimum et son maximum sur $[a, b]$, c'est-à-dire qu'il existe des points $\alpha, \beta \in [a, b]$ tels que $f(\alpha) \leqslant f(x) \leqslant f(\beta)$ pour tous les $x \in [a, b]$. Autrement dit,*

$$\min_{x \in [a, b]} f(x) = f(\alpha), \qquad \max_{x \in [a, b]} f(x) = f(\beta).$$

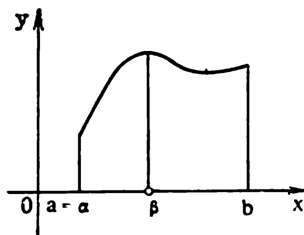


Fig. 26

La fonction continue $y = f(x)$ représentée sur la figure 25 atteint son minimum sur $[a, b]$ au point $x = \alpha$ et son maximum au point $x = \beta$. Ici les points $\alpha$ et $\beta$ appartiennent à l'intervalle $]a, b[$. La fonction continue $y = f(x)$ dont le graphe est représenté sur la figure 26 atteint son minimum en l'origine de l'intervalle $[a, b]$ et son maximum en un point intérieur $\beta$ de $[a, b]$.

REMARQUE 1. Le théorème 1 nous dit que la fonction $f$ est bornée sur $[a, b]$, car continue. Donc, $\inf f(x)$ et $\sup f(x)$ existent dans $[a, b]$ et

$$\inf_{x \in [a, b]} f(x) \leqslant \sup_{x \in [a, b]} f(x).$$

Le théorème 2 affirme que ces bornes sont atteintes sur $[a, b]$, c'est-à-dire qu'on peut remplacer inf et sup respectivement par min et max.

REMARQUE 2. La fonction $y = x$ est continue et bornée sur l'intervalle $]0, 1[$ ; elle n'atteint pas sa borne supérieure $\sup\limits_{x \in ]0,1[} x = 1$, c'est-à-dire qu'il n'existe pas un $x_0 \in ]0, 1[$ tel que $y = 1$. Donc, dans le théorème 2, la condition de continuité de $f$ sur un *intervalle fermé* borné est essentielle.

Il est évident que $\sup\limits_{x \geqslant 0} \operatorname{Arctg} x = \pi/2$. Cependant, il n'existe pas un $x \geqslant 0$ tel que $\operatorname{Arctg} x = \pi/2$. Les hypothèses du théorème sont mises en défaut, car le domaine de définition de la fonction continue $\operatorname{Arctg} x$ est illimité.

Si une fonction $f$ est discontinue sur $[a, b]$, elle n'atteint pas forcément son suprémum. Exemple, la fonction

$$f(x) = \begin{cases} x, & 0 \leqslant x < 1/2, \\ 0, & 1/2 \leqslant x \leqslant 1. \end{cases}$$

THÉORÈME 3. *Si une fonction $f$ est continue sur un intervalle $[a, b]$, et $f(a)$ et $f(b)$ sont différents de 0 et de signes contraires, alors il existe au moins un point $c \in ]a, b[$ en lequel $f(c) = 0$.*

La fonction dont le graphe $\Gamma$ est représenté sur la figure 27 remplit les conditions du théorème 3. Elle est continue sur $[a, b]$, et $f(a) < 0$, $f(b) > 0$. Il est évident pour des raisons géométriques que le graphe $\Gamma$ doit couper l'axe $Ox$ au moins en un point $c \in ]a, b[$.

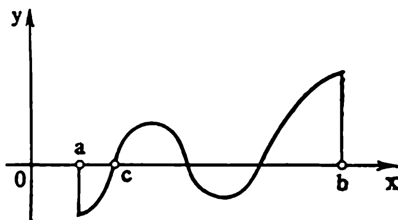


Fig. 27

COROLLAIRE 1. *Si une fonction $f$ est continue sur $[a, b]$, $f(a) = A$, $f(b) = B$ $(A \neq B)$ et $C$ est un nombre compris entre $A$ et $B$, alors il existe au moins un point $c \in ]a, b[$ tel que $f(c) = C$.*

Ce corollaire s'énonce encore: *une fonction continue sur un intervalle $[a, b]$ prend toutes les valeurs intermédiaires entre ses valeurs aux extrémités de l'intervalle $[a, b]$.*

COROLLAIRE 2. *Une fonction continue sur un intervalle $[a, b]$ prend toutes les valeurs intermédiaires entre son infimum et son suprémum.*

DÉMONSTRATION DU COROLLAIRE 1. Définissons une nouvelle fonction $F(x) = f(x) - C$, où $C$ est une constante comprise entre $A$ et $B$. La fonction $f$ étant continue sur $[a, b]$, il en est de même de la fonction $F$. De plus, $F$ prend de toute évidence des valeurs de signes contraires aux extrémités de $[a, b]$. Donc, le théorème 3 affirme

l'existence dans $]a, b[$ d'un point $c$ tel que $F(c) = 0$ ou $f(c) - C = 0$, d'où $f(c) = C$, c. q. f. d.

EXEMPLE. L'équation $x - \cos x = 0$ présente une racine dans l'intervalle $]0, \pi[$.

En effet, la fonction $f(x) = x - \cos x$ est continue sur l'intervalle $[0, \pi]$ et prend aux extrémités de celui-ci des valeurs de signes contraires: $f(0) = -1$, $f(\pi) = \pi + 1$.

Les théorèmes 1, 2 et 3 seront démontrés formellement plus bas.

DEMONSTRATION DU THEOREME 1. Supposons que $f$ n'est pas bornée sur $[a, b]$. Pour tout entier naturel $n$ il existe alors un point $x_n \in [a, b]$ tel que

$$|f(x_n)| > n \quad (n = 1, 2, \ldots). \tag{1}$$

La suite $\{x_n\}$ est bornée (car $a$ et $b$ sont finis) et on peut en extraire une suite $\{x_{n_k}\}$ convergeant vers un nombre $\alpha \in [a, b]$ (voir corollaire du théorème 4, § 1, chap. 2). Or, la fonction $f$ est continue en $\alpha$ (si $\alpha = a$ (resp. $\alpha = b$), la fonction $f$ est continue à droite (resp. à gauche)), donc

$$\lim_{k \to \infty} f(x_{n_k}) = f(\alpha). \tag{2}$$

La relation (2) contredit la relation (1), donc $f$ ne peut être que bornée sur $[a, b]$.

DEMONSTRATION DU THEOREME 2. Le théorème précédent dit qu'une fonction continue sur $[a, b]$ est bornée, c'est-à-dire que

$$f(x) \leqslant K \quad (x \in [a, b]).$$

La fonction $f$ admet donc une borne supérieure sur $[a, b]$:

$$\sup_{x \in [a, b]} f(x) = M. \tag{3}$$

Le nombre $M$ jouit de la propriété suivante: pour tout entier naturel $n$ il existe dans $[a, b]$ un point $x_n$ tel que

$$M - \frac{1}{n} < f(x_n) \leqslant M \quad (n = 1, 2, \ldots).$$

La suite $\{x_n\}$ est bornée, car appartenant à $[a, b]$, donc on peut en extraire une suite $\{x_{n_k}\}$ convergeant vers un nombre $\beta \in [a, b]$. Or, la fonction $f$ est continue en $\beta$, donc

$$\lim_{k \to \infty} f(x_{n_k}) = f(\beta).$$

D'autre part, $M - \frac{1}{n_k} < f(x_{n_k}) \leqslant M$ $(k = 1, 2, \ldots)$ et $\lim\limits_{k \to \infty} f(x_{n_k}) = M$.

Mais comme $f(x_{n_k})$ ne peut tendre que vers une seule limite, on a $M = f(\beta)$. Donc, *la fonction $f$ atteint son maximum en un point*

$\beta \in [a, b]$. Nous avons prouvé qu'il existait un point $\beta \in [a, b]$ tel que

$$\max_{x \in [a, b]} f(x) = f(\beta).$$

La partie du théorème relative au minimum se démontre de façon analogue. Cependant, il est possible de ramener cette démonstration à celle de la première partie en tenant compte du fait que

$$\min_{x \in [a, b]} f(x) = -\max_{x \in [a, b]} \{-f(x)\}.$$

DÉMONSTRATION DU THÉORÈME 3. Désignons l'intervalle $[a, b]$ par $\sigma_0$. Divisons $\sigma_0$ en deux parties égales. Si la fonction s'annule au milieu de $\sigma_0$, le théorème est démontré ; sinon, l'une des deux parties de $\sigma_0$ est telle que la fonction $f$ prend en ses extrémités des valeurs de signes contraires. Désignons cette partie par $\sigma_1$ et divisons-la en deux parties égales. Si la fonction s'annule au milieu de $\sigma_1$, le théorème est démontré ; sinon, désignons par $\sigma_2$ celle des deux parties de $\sigma_1$ aux extrémités de laquelle la fonction prend des valeurs de signes contraires. En poursuivant cette procédure, soit on trouve un point $c \in ]a, b[$ tel que $f(c) = 0$ et le théorème sera prouvé, soit on obtient une suite illimitée d'intervalles emboîtés $\sigma_0 \supset \sigma_1 \supset \sigma_2 \supset \ldots$ aux extrémités de chacun desquels la fonction prend des valeurs de signes contraires. Il existe alors un point $c$ appartenant à tous les $\sigma_n$, donc à $[a, b]$. De toute évidence, $f(c) = 0$, car si l'on admet par exemple que $f(c) > 0$ il existerait alors un voisinage $V(c)$ du point $c$ tel que pour tous les $x \in V(c)$ la fonction $f(x)$ serait positive. Or, ceci est impossible, car pour $n$ assez grand $\sigma_n \subset V(c)$, et $f$ change de signe sur $\sigma_n$. Ce qui prouve le théorème.

## § 6. Fonction réciproque continue

Considérons une fonction continue $y = f(x)$ strictement croissante sur $[a, b]$ (fig. 28). Soit

$$f(a) = \alpha, \quad f(b) = \beta.$$

Le graphe de cette fonction est une courbe continue. L'on voit sur ce graphe que si $x$ croît continûment de $a$ à $b$, alors $y$ croît continûment de $\alpha$ à $\beta$ en parcourant une fois toutes les valeurs de l'intervalle $[\alpha, \beta]$. Mais alors à chaque valeur $y$ correspond une valeur unique $x \in [a, b]$ telle que $y = f(x)$. Ceci définit sur $[\alpha, \beta]$ une fonction

$$x = g(y),$$

appelée *fonction réciproque* de la fonction $y = f(x)$.

Il est évident que la fonction $x = g(y)$ est strictement croissante sur l'intervalle $[\alpha, \beta]$ et applique cet intervalle sur l'intervalle

$[a, b]$ : on a les identités

$$f[g(y)] = y, \quad \forall y \in [\alpha, \beta],$$
$$g[f(x)] = x, \quad \forall x \in [a, b].$$

On déduit le graphe de la fonction $x = g(y)$ à partir de celui de la fonction $f(x)$ par une rotation de $2\pi$ autour de la bissectrice du premier quadrant. Comme la rotation ne modifie pas la continuité du graphe, la fonction $x = g(y)$ est continue sur $[\alpha, \beta]$.

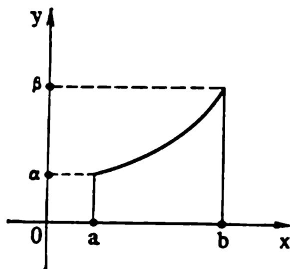



Fig. 28

Donc, on a établi, par des considérations purement géométriques, le théorème suivant.

THÉORÈME 1. *Supposons qu'une fonction $f$ est continue et strictement croissante sur un intervalle $[a, b]$ et que $\alpha = f(a)$, $\beta = f(b)$. Alors*: 1) *l'image de l'intervalle $[a, b]$ par $f$ est l'intervalle $[\alpha, \beta]$,* 2) *il existe une fonction $x = g(y)$ réciproque de $f$, strictement croissante et continue sur $[\alpha, \beta]$.*

La démonstration formelle de ce théorème repose sur le lemme suivant.

LEMME 1. *Supposons qu'une fonction strictement croissante $y = f(x)$ applique un intervalle $[a, b]$ sur un intervalle $[\alpha, \beta]$, c'est-à-dire que $f([a, b]) = [\alpha, \beta]$. Alors $f$ est continue sur $[a, b]$.*

DÉMONSTRATION. Considérons un point quelconque $x_0 \in ]a, b[$ Comme la fonction $f$ est strictement croissante, le point correspondant $y_0 = f(x_0)$ appartiendra à l'intervalle $]\alpha, \beta[$.

Prenons un $\varepsilon > 0$ assez petit pour que $\alpha < y_0 - \varepsilon < y_0 < y_0 + \varepsilon < \beta$. Par hypothèse, on peut exhiber des points $x_1$ et $x_2$ de $]a, b[$ $(x_1 < x_0 < x_2)$ tels que $y_0 - \varepsilon = f(x_1)$, $y_0 + \varepsilon = f(x_2)$.

L'intervalle $]x_1, x_2[$ peut être traité comme un voisinage du point $x_0$.

La fonction $f$ étant strictement croissante, pour $x \in ]x_1, x_2[$ on aura $y_0 - \varepsilon < f(x) < y_0 + \varepsilon$ ou $|f(x) - y_0| < \varepsilon$, c'est-à-dire que

$$|f(x) - f(x_0)| < \varepsilon,$$

ce qui exprime que la fonction $f(x)$ est continue en $x_0$.

Si $x_0 = a$ ou $x_0 = b$, on démontre par analogie la continuité à gauche et à droite de la fonction $f$.

DÉMONSTRATION DU THÉORÈME 1. Supposons que $Y = f([a, b])$ est l'image de $[a, b]$ par $f$. Comme $\alpha = f(a)$, $\beta = f(b)$ et $f$ est strictement croissante, on a $\alpha \leqslant f(x) \leqslant \beta$, $\forall x \in [a, b]$, d'où il s'ensuit que

$$Y \subset [\alpha, \beta]. \tag{1}$$

D'autre part, si $f$ est un point quelconque de l'intervalle $[\alpha, \beta]$, il appartient à $Y$ en vertu du théorème 3 du § 5 des valeurs intermédiaires d'une fonction continue, c'est-à-dire que

$$[\alpha, \beta] \subset Y. \tag{2}$$

De (1) et de (2) on déduit la proposition 1):

$$Y = [\alpha, \beta].$$

La proposition 2) dérive maintenant du lemme 1. En effet, la fonction $f$ étant strictement croissante, elle admet sur $Y = [\alpha, \beta]$ une fonction réciproque $g(y)$ strictement croissante qui applique l'intervalle $[\alpha, \beta]$ sur $[a, b]$. Donc, cette fonction est continue d'après le lemme 1. Ceci prouve le théorème.

En modifiant légèrement les raisonnements ci-dessus, on démontre l'analogue suivant du théorème 1.

Théorème 1′. *Supposons qu'une fonction $f$ est continue et strictement croissante sur $]a, b[$ (ou sur $[a, b[$ ou encore sur $]a, b]$) et que*

$$\alpha = \inf_{x\in]a,\, b[} f(x), \quad \beta = \sup_{x\in]a,\, b[} f(x).$$

*Alors l'image de l'intervalle $]a, b[$* (resp. $[a, b[$, $]a, b]$) *est l'intervalle $]\alpha, \beta[$ (resp. $[\alpha, \beta[$, $]\alpha, \beta]$) et la fonction $x = g(y)$ réciproque de $f$ est univalente, strictement croissante et continue sur $]\alpha, \beta[$ (resp. $[\alpha, \beta[$, $]\alpha, \beta]$).*

Remarque. La réciproque d'une fonction $f(x)$ strictement décroissante et continue sur $[a, b]$ (resp. $]a, b[$) est une fonction continue strictement décroissante sur $[\beta, \alpha]$, où $\alpha = f(a)$, $\beta = f(b)$. On établit ceci sans peine en considérant la fonction $-f(x)$ ou la fonction $f(-x)$.

Si une fonction $f(x)$ est continue mais non strictement monotone sur $[a, b]$, sa réciproque est une correspondance qui à certains $y$ associe plusieurs valeurs de $x$.

Exemple. La fonction

$$y = \sin x, \quad x \in ]-\infty, \infty[,$$

est continue mais pas monotone. Son domaine de valeurs est l'intervalle $[-1, 1]$. A chaque $y$ de cet intervalle correspond une infinité de valeurs de $x$ telles que $y = \sin x$.

Du reste, sur l'intervalle $[-\pi/2, \pi/2]$ la fonction $y = \sin x$ est continue et strictement croissante, et sa réciproque est une fonction continue qui, nous le savons, se désigne par:

$$x = \operatorname{Arcsin} y, \quad y \in [-1, 1].$$

## § 7. Continuité uniforme d'une fonction

Soit une fonction $f$ continue sur un intervalle fermé $[a, b]$ (resp. ouvert ou semi-ouvert). Alors pour chaque point $x_0$ de cet intervalle on peut exhiber pour un $\varepsilon > 0$ donné un $\delta > 0$ tel que

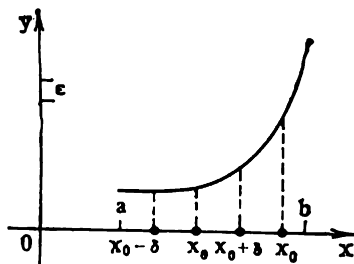


Fig. 29

$$|f(x) - f(x_0)| < \varepsilon,$$

dès que

$$|x - x_0| < \delta, \quad x \in [a, b]$$

$$(\text{resp. } ]a, b[, \ [a, b[, \ ]a, b]).$$

Le nombre $\delta$ varie en général avec $x_0$, à $\varepsilon$ constant: il dépend en effet de $\varepsilon$ et de $x_0$. Sur la figure 29 on voit que le nombre $\delta$ peut convenir pour une partie en pente douce du graphe et être trop grand pour une partie raide.

Il semble donc naturel de distinguer les fonctions continues pour lesquelles on peut exhiber pour un $\varepsilon > 0$ un $\delta > 0$ qui convienne à la fois pour tous les $x$ de l'ensemble de définition.

Commençons par une définition.

DEFINITION. *On dit qu'une fonction $f$ définie sur un ensemble $X$ est uniformément continue sur $X$ si pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ dépendant uniquement de $\varepsilon$, tel que*

$$|f(x') - f(x'')| < \varepsilon$$

*pour tous les $x'$, $x'' \in X$ tels que $|x' - x''| < \delta$.*

Il est aisé de voir que si une fonction est uniformément continue sur un ensemble $X$, elle l'est à fortiori sur l'une quelconque des parties de $X$. La réciproque n'est pas toujours vraie.

THEOREME. *Si une fonction $f$ est définie et continue sur $[a, b]$, alors elle est uniformément continue sur $[a, b]$.*

DEMONSTRATION. Supposons par absurde que le théorème est faux. Il existe alors un $\varepsilon > 0$ tel que pour tout $\delta > 0$ on peut exhiber un couple de points $x'$ et $x''$ de $[a, b]$ tels que $|x' - x''| < \delta$ et que

$$|f(x') - f(x'')| \geqslant \varepsilon.$$

Considérons une suite de nombres strictement positifs $\delta_n$ ($n = 1, 2, \ldots$) convergeant vers 0. Pour chaque $\delta_n$ il existe des points $x'_n$, $x''_n \in [a, b]$ tels que

$$|x'_n - x''_n| < \delta_n, \quad \text{mais } |f(x'_n) - f(x''_n)| \geqslant \varepsilon. \tag{1}$$

La suite $\{x'_n\}$ est bornée, car ses points appartiennent à $[a, b]$. Le théorème de Bolzano-Weierstrass nous dit qu'on peut en extraire une suite $\{x'_{n_k}\}$ convergeant vers un point $x_0 \in [a, b]$. La suite extraite $\{x''_{n_k}\}$ converge aussi vers 0, puisque $x'_{n_k} - x''_{n_k} \to 0$, $k \to \infty$. Par hypothèse, la fonction $f$ est continue sur $[a, b]$, donc en $x_0$. Si $x_0 = a$ (resp. $x_0 = b$), on admettra que $f$ est continue en $x_0$ à droite (resp. à gauche). Donc

$$\lim_{k\to\infty} f(x'_{n_k}) = \lim_{k\to\infty} f(x''_{n_k}) = f(x_0).$$

Dans (1), en passant à la limite pour $k \to \infty$, on obtient

$$\varepsilon \leqslant \lim_{k\to\infty} |f(x'_{n_k}) - f(x''_{n_k})| = |f(x_0) - f(x_0)| = 0, \qquad (2)$$

ce qui est absurde, puisque $\varepsilon > 0$.

Signalons que dans (2) on s'est servi de la continuité de la fonction $|u|$ (cf. § 3, exemple 8). Ceci achève la démonstration du théorème.

EXEMPLE. La fonction

$$y = \sin(1/x)$$

est continue sur l'intervalle $[\delta, 1]$, $\forall \delta > 0$, donc elle est uniformément continue sur cet intervalle en vertu du théorème ci-dessus.

On remarquera que cette fonction est continue sur l'intervalle $]0, 1]$ mais pas uniformément. Ceci montre qu'il est essentiel que dans le théorème la fonction soit donnée sur un intervalle fermé et non sur un intervalle semi-ouvert ou ouvert.

Assurons-nous que la fonction $y = \sin(1/x)$ n'est pas uniformément continue sur l'intervalle $]0, 1]$. Les points $x_k = \frac{2}{\pi(2k+1)}$ ($k = 0, 1, 2, \ldots$) appartiennent visiblement à l'intervalle $]0, 1]$ et

$$|f(x_{k+1}) - f(x_k)| = \left|\sin\frac{\pi(2k+3)}{2} - \sin\frac{\pi(2k+1)}{2}\right| = \\ = |(-1)^{k+1} - (-1)^k| = 2.$$

Si l'on prend $\varepsilon = 1$, alors pour tout $\delta > 0$ il existe un $k$ tel que

$$|x_{k+1} - x_k| = \frac{4}{\pi(2k+3)(2k+1)} < \delta,$$

alors que

$$|f(x_{k+1}) - f(x_k)| = 2 > \varepsilon = 1.$$

De ce qui vient d'être dit il s'ensuit que la fonction $y = \sin(1/x)$ ne peut être prolongée en continuité à l'intervalle fermé $[0, 1]$, car, en vertu du théorème démontré, elle serait uniformément continue sur cet intervalle, donc sur l'intervalle $]0, 1]$, ce qui est impossible.

## § 8. Fonctions élémentaires

On appelle ainsi les fonctions $C$ (constante), $x^n$, $a^x$, $\log_a x$, $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{Arcsin} x$, $\operatorname{Arccos} x$, $\operatorname{Arctg} x$.

En appliquant à ces fonctions les opérations arithmétiques et l'opération de superposition un nombre fini de fois, on obtiendrait des fonctions plus complexes que nous appellerons aussi *fonctions élémentaires.*

Par exemple, $y = \ln(e^x + \sin^2 x + 1)$ est une fonction élémentaire.

Etudions les propriétés de ces fonctions.

a) Fonction constante $C$. Cette fonction associe à chaque nombre réel le nombre $C$. Son graphe est une droite parallèle à l'axe $Ox$ et située à une distance $|C|$ au-dessus de $Ox$ si $C > 0$ et au-dessous si $C < 0$. Cette fonction est continue sur l'axe des $x$ tout entier (voir § 3, exemple 1).

b) Fonction puissance $x^n$ ($n$ est une constante). Pour $n \in \mathbf{N}$, cette fonction est définie sur l'axe des $x$ tout entier. Pour la calculer (théoriquement!) on représente $x$ par sa forme décimale ($x = \pm \alpha_0, \alpha_1, \alpha_2 \ldots$) et on multiplie cette forme $n$ fois par elle-même suivant la règle de multiplication des formes décimales (cf. § 6, (11), chap. 1) et la règle des signes.

La fonction $x^n$ est continue en tant que produit fini de fonctions continues ($y = x$). Elle est strictement croissante sur $[0, \infty[$, puisque

$$x_2^n - x_1^n = (x_2 - x_1)(x_2^{n-1} + x_2^{n-2}x_1 + \ldots + x_1^{n-1}) > 0$$

pour $x_1 < x_2$. De plus, elle tend vers $+\infty$ avec $x$. En effet, si $x \geqslant 1$, alors $x^n = x^{n-1}x \geqslant x$ ($n > 1$) et $x^n \to \infty$ avec $x$.

Ainsi, la fonction $\varphi(x) = x^n$, pour $n \in \mathbf{N}$, est continue, strictement croissante sur $[0, \infty[$ et telle que $\varphi(0) = 0$, $\sup\limits_{x \in [0, \infty[} \varphi(x) = +\infty$. Donc, le théorème 1′ du § 6 nous dit que la fonction $y = x^n$ applique l'intervalle semi-ouvert $X = [0, \infty[$ sur l'intervalle semi-ouvert $Y = [0, \infty[$ et admet une fonction réciproque continue, strictement croissante. Cette fonction se note

$$x = y^{1/n} = \sqrt[n]{y} \quad (y \geqslant 0)$$

et s'appelle *valeur arithmétique de la racine n-ième de y.*

Signalons que pour $y > 1$ ($n \geqslant 1$)

$$\sqrt[n]{y} > \sqrt[n]{1} = 1. \tag{1}$$

Notons que *si a est un nombre positif arbitraire* ($a \in [0, \infty[$), *d'après ce qui précède on peut exhiber un nombre positif unique* $b = a^{1/n}$ *tel que* $b^n = a$.

Nous venons de *prouver l'existence de la racine n-ième de* $a \geqslant 0$.

Observons que si $n$ est impair, c'est-à-dire $n = 2k + 1$ ($k = 0$, 1, 2, . . .), la fonction $y = x^n$ est impaire ($(-x)^n = -x^n$). Elle est continue, strictement croissante sur $]-\infty, \infty[$ et telle que

$$\inf_{x \in ]-\infty, \infty[} x^n = -\infty, \quad \sup_{x \in ]-\infty, \infty[} x^n = +\infty.$$

Donc, en vertu du théorème 1′ du § 6, la fonction $y = x^{2k+1}$ applique l'intervalle $]-\infty, \infty[$ sur lui-même et admet sur $]-\infty, \infty[$ une

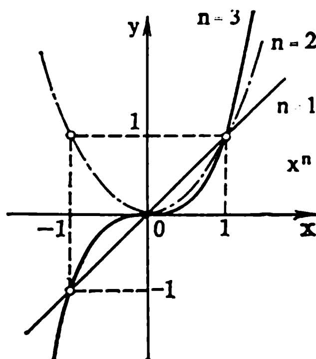



Fig. 30

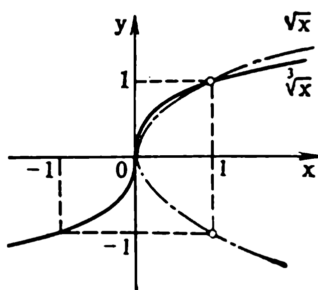



Fig. 31

fonction réciproque continue, strictement croissante

$$x = \sqrt[n]{y}, \quad y \in ]-\infty, \infty[, \quad n = 2k + 1.$$

L'expression $\sqrt[n]{y}$, avec $y > 0$, désigne ici la valeur arithmétique de la racine $n$-ième de $y$, c'est-à-dire le nombre positif dont la puissance $n$-ième est $y$. Si $y < 0$, alors

$$\sqrt[n]{y} = -\sqrt[n]{|y|}.$$

Pour $n = 2k$ ($k = 1, 2, \ldots$) la fonction $y = x^n$ est paire et continue. Elle applique l'intervalle $]-\infty, \infty[$ sur l'intervalle semi-ouvert $[0, \infty[$. Elle n'est pas monotone sur $]-\infty, \infty[$ et sa réciproque est une correspondance bivalente

$$x = \pm \sqrt[n]{y} \quad (y \geqslant 0).$$

La réciproque ne sera une fonction que pour $y = 0$.

Les graphes des figures 30 et 31 représentent les fonctions $x^n$ et $x^{1/n}$ pour différents $n$.

La fonction $x^n$ se définit aussi pour $n$ rationnel. Soient $p, q \in \mathbb{N}$. On pose

$$x^{p/q} = \sqrt[q]{x^p} \quad (x \geqslant 0)$$

et on démontre que

$$x^{p/q} = x^{sp/sq} = (\sqrt[q]{x})^p \quad (x \geqslant 0).$$

On pose aussi

$$x^{-p/q} = \frac{1}{\sqrt[q]{x^p}} \quad (x > 0),$$
$$x^0 = 1.$$

Ceci étant, on démontre pour tout $n$ rationnel la propriété caractéristique de la fonction puissance:

$$(xy)^n = x^n y^n \quad (x, y > 0).$$

On établit sans peine que la fonction $y = x^{p/q}$ $(p, q \in \mathbf{N})$ est continue et strictement croissante sur l'intervalle $[0, \infty[$, et applique l'intervalle $[0, \infty[$ sur lui-même, donc elle admet une inverse continue, strictement croissante, définie visiblement par

$$x = y^{q/p}, \quad y \in [0, \infty[.$$

S'agissant de la fonction $y = x^{-p/q}$ $(p, q \in \mathbf{N})$, elle est strictement décroissante, continue sur l'intervalle $]0, \infty[$ et applique l'intervalle $]0, \infty[$ sur lui-même. Sa réciproque sur $]0, \infty[$ est donc une fonction continue, strictement décroissante, définie par

$$x = y^{-q/p} \quad (y > 0).$$

Il est évident que

$$\lim_{\substack{x \to 0 \\ x > 0}} x^{-p/q} = +\infty, \quad \lim_{\substack{y \to 0 \\ y > 0}} y^{-q/p} = +\infty$$

On peut définir la fonction puissance $x^n$ (pour $x > 0$, et pour $x = 0$ si $n \geqslant 0$) pour $n$ irrationnel, mais il vaut mieux le faire à l'aide de la fonction exponentielle $a^x$ (cf. c) plus bas).

On ne s'est intéressé qu'aux racines réelles de l'équation $y = x^n$. Si l'on avait cherché les racines complexes, on aurait trouvé $n$ racines distinctes pour chaque $y \neq 0$ (cf. § 3, chap. 5).

Exemple. Montrons que

$$\lim_{n \to \infty} \sqrt[n]{n} = 1. \tag{2}$$

En effet, la formule du binôme de Newton nous donne pour $\lambda > 0$:

$$(1+\lambda)^n = 1 + n\lambda + \frac{n(n-1)}{2!}\lambda^2 + \ldots > 1 + \frac{n(n-1)}{2}\lambda^2.$$

Si l'on admet que $\lambda = \sqrt[n]{n} - 1$ ($>0$ pour $n \geqslant 2$, voir (1)), on obtient

$$n > 1 + \frac{n(n-1)}{2}(\sqrt[n]{n} - 1)^2$$

ou

$$\frac{2}{n} > (\sqrt[n]{n}-1)^2 > 0 \quad (n \geqslant 2).$$

La fonction $\sqrt{x}$ étant strictement monotone, on obtient

$$\sqrt{2/n} > \sqrt[n]{n}-1 > 0$$

ou

$$\sqrt{2/n}+1 > \sqrt[n]{n} > 1 \quad (n \geqslant 2). \tag{3}$$

Enfin, la fonction $\sqrt{x}$ étant continue en $x = 0$, on obtient la relation (2) en passant à la limite pour $n \to \infty$ dans (3) (cf. théorème 5 du § 1, chap. 2).

c) F o n c t i o n e x p o n e n t i e l l e $a^x (a > 0, a \neq 1)$. Dans la suite on désignera les nombres rationnels par les lettres grecques $\alpha$, $\beta$, $\gamma$, . . . Soit $a > 1$.

On sait du cours du secondaire ce que c'est que $a^x$ pour $x$ rationnel (voir encore b)).

On sait aussi que:

$1_1$) $a^\alpha > 0$,

$2_1$) $a^\alpha a^\beta = a^{\alpha+\beta}$,

$3_1$) $a^\alpha < a^\beta$ ($\alpha < \beta$, $a > 1$),

$4_1$) $(a^\alpha)^\beta = a^{\alpha\beta}$,

$5_1$) $a^{\alpha_n} \to +\infty$, $\alpha_n \to +\infty$, $a > 1$.

Prouvons la propriété $5_1$). Mettons $a$ sous la forme $a = 1 + \lambda$ ($\lambda > 0$!). Alors

$$a^n = (1+\lambda)^n = 1 + n\lambda + \ldots > 1 + n\lambda.$$

Le second membre de cette inégalité tend vers l'infini avec $n$, donc il en est de même du premier. Si l'on désigne par $[\alpha_n]$ la partie entière de $\alpha_n$, on aura

$$a^{\alpha_n} \geqslant a^{[\alpha_n]},$$

et si $\alpha_n \to +\infty$, alors $[\alpha_n] \to +\infty$, donc $a^{[\alpha_n]} \to +\infty$. Alors $a^{\alpha_n} \to +\infty$.

Soit $x$ un nombre rationnel quelconque. Montrons qu'*on peut définir le nombre $a^x$ comme la borne supérieure des nombres $a^\alpha$ étendue à tous les nombres rationnels $\alpha < x$*:

$$\sup_{\alpha < x} a^\alpha = a^x. \tag{4}$$

En effet, d'après $3_1$) on a

$$a^\alpha < a^x, \quad \forall \alpha < x, \tag{5}$$

c'est-à-dire la première propriété de la borne supérieure. Construisons une suite strictement croissante de nombres rationnels $\alpha_n$, convergeant vers $x$. On a pour cette suite : $a^{\alpha_n} \to a^x$, $n \to \infty$. On prouve cette propriété plus bas (voir (8)). Donc, pour tout $\varepsilon > 0$ on peut exhiber un $n$ tel que

$$a^x - \varepsilon < a^{\alpha_n} < a^x, \tag{6}$$

c'est-à-dire qu'on obtient la deuxième propriété de la borne supérieure. De (5) et (6) on déduit (4).

Soient maintenant un nombre irrationnel $x$ et un entier naturel $m$ supérieur à $x$ ($m > x$). On a l'inégalité évidente

$$\underset{\alpha<x}{a^\alpha} < a^m,$$

c'est-à-dire que l'ensemble des nombres $a^\alpha$ tel que $\alpha < x$ est majoré. Il admet donc une borne supérieure

$$\sup_{\alpha<x} a^\alpha.$$

Cette borne est un nombre bien déterminé que nous désignerons par $a^x$:

$$a^x = \sup_{\alpha<x} a^\alpha. \tag{7}$$

Ceci définit la fonction $a^x$ pour tous les $x$ réels. On l'appelle *fonction exponentielle.*

Donc, *la fonction $a^x$ se définit comme la borne supérieure des nombres $a^\alpha$, étendue à tous les nombres rationnels $\alpha < x$.*

Si $x$ est un nombre rationnel, cette définition coïncide avec l'ancienne (elle donne le même nombre); si $x$ est irrationnel, elle donne de nouveaux nombres $a^x$.

On démontre que la fonction $a^x$ définie à l'aide de (7) possède les propriétés importantes suivantes:

1) $a^x > 0$,
2) $a^x a^y = a^{x+y}$,
3) $a^x < a^y \quad (x < y,\ a > 1)$,
4) $a^x \to 1, \quad x \to 0$,
5) $a^x \to \infty, \quad x \to \infty \ (a > 1)$
6) $a^x \to 0, \quad x \to -\infty \ (a > 1)$,
7) $a^{xy} = (a^x)^y$.

Signalons que les propriétés 2) et 4) entraînent la continuité de la fonction $a^x$ pour tout $x_0 \in ]-\infty, \infty[$:

$$|a^x - a^{x_0}| = |a^{x_0}(a^{x-x_0} - 1)| = a^{x_0}|a^{x-x_0} - 1| \to 0 \tag{8}$$

pour $x - x_0 \to 0$.

Dans la suite on admettra que $a > 1$.

DÉMONSTRATION DE 1). Pour tout $x$ il existe $\alpha_0 < x$, donc

$$0 < a^{\alpha_0} < \sup_{\alpha < x} a^{\alpha} = a^x, \quad \text{i.e. } 0 < a^x.$$

DÉMONSTRATION DE 2). Soient $\alpha < x$ et $\beta < y$, alors $\alpha + \beta < x + y$. Soit par ailleurs $\gamma$ un nombre rationnel tel que

$$\gamma < x + y. \tag{9}$$

Montrons que $\gamma$ peut se mettre sous la forme

$$\gamma = \alpha + \beta, \quad \text{où} \quad \alpha < x, \quad \beta < y. \tag{10}$$

De (9) il résulte que

$$\gamma - y < x.$$

Choisissons un nombre rationnel $\alpha$ tel que

$$\gamma - y < \alpha < x, \tag{11}$$

et posons

$$\beta = \gamma - \alpha. \tag{12}$$

De la première inégalité (11) il s'ensuit alors

$$\beta < y. \tag{13}$$

Donc, l'ensemble de toutes les sommes $\alpha + \beta$, où $\alpha < x$, $\beta < y$, est égal à celui de tous les $\gamma < x + y$:

$$\{\alpha+\beta\}_{\substack{\alpha<x\\ \beta<y}} = \{\gamma\}_{\gamma<x+y}.$$

Donc

$$a^{x+y} = \sup_{\gamma<x+y} a^{\gamma} = \sup_{\substack{\alpha<x\\ \beta<y}} a^{\alpha+\beta} = \sup_{\substack{\alpha<x\\ \beta<y}} (a^{\alpha}a^{\beta}) = \sup_{\alpha<x} a^{\alpha} \sup_{\beta<y} a^{\beta} = a^x \cdot a^y$$

(cf. chap. 2, § 8, exercice 2).

DÉMONSTRATION DE 3). Soient $x < y$ et $\beta_1$ et $\beta_2$ des nombres rationnels tels que $x < \beta_1 < \beta_2 < y$. Alors

$$a^x = \sup_{\alpha<x} a^{\alpha} \leqslant \sup_{\alpha<\beta_1} a^{\alpha} = a^{\beta_1} < a^{\beta_2} \leqslant \sup_{\beta<y} a^{\beta} = a^y,$$

et nous avons prouvé que

$$a^x < a^y.$$

DÉMONSTRATION DE 4). Pour $n > a$ ($n \in N$, $a > 1$!)

$$1 < a^{1/n} < n^{1/n} \xrightarrow[n\to\infty]{} 1$$

(voir exemple de b)), et pour l'instant on a démontré que

$$\lim_{n\to\infty} a^{1/n} = 1. \tag{14}$$

Considérons maintenant une suite quelconque de nombres strictement positifs $x_n < 1$ convergeant vers 0. Soit $k_n = [1/x_n]$. Alors $0 < x_n < 1/k_n$ et

$$1 = a^0 < a^{x_n} \leqslant a^{1/k_n}.$$

Donc, en vertu de (14)

$$\lim_{n\to\infty} a^{x_n} = 1.$$

La suite $\{x_n\}$ étant arbitraire, on a prouvé l'existence de la limite à droite

$$\lim_{\substack{x\to 0\\ x>0}} a^x = 1, \tag{15}$$

et par suite celle de la limite à gauche

$$\lim_{\substack{x\to 0\\ x<0}} a^x = \lim_{\substack{-x\to 0\\ -x>0}} \frac{1}{a^{-x}} = \frac{1}{\lim\limits_{\substack{u\to 0\\ u>0}} a^u} = \frac{1}{1} = 1. \tag{16}$$

De (15) et de (16) il résulte que

$$\lim_{x\to 0} a^x = 1$$

(voir § 2, (6), (7)). Ceci achève la démonstration de 4).

DÉMONSTRATION DE 5). Soit un nombre $M > 0$ arbitrairement grand. Il existe un nombre rationnel $\alpha$ tel que $a^\alpha > M$, donc

$$M < a^\alpha < a^x, \quad \forall x > \alpha.$$

Donc, $a^x \to +\infty$ avec $x$.

DÉMONSTRATION DE 6)

$$\lim_{x\to -\infty} a^x = \lim_{-x\to +\infty} \frac{1}{a^{-x}} = \frac{1}{\lim\limits_{u\to +\infty} a^u} = 0.$$

DÉMONSTRATION DE 7). Notons que pour $m$ naturel, on a en vertu de 2)

$$a^{xm} = a^x \ldots a^x = (a^x)^m; \quad \left(a^{\frac{x}{m}}\right)^m = a^{\frac{x}{m}} \ldots a^{\frac{x}{m}} =$$

$$= a^{\frac{x}{m}\cdot m} = a^x; \qquad a^{\frac{x}{m}} = (a^x)^{\frac{1}{m}}.$$

Pour le nombre rationnel $\frac{p}{q} > 0$

$$(a^x)^{\frac{p}{q}} = \left(a^x\right)^{\frac{1}{q}\cdot p} = \left(a^{\frac{x}{q}}\right)^p = a^{x\frac{p}{q}}.$$

Par ailleurs, si $y$ est un nombre strictement positif arbitraire et $\alpha_n \to y$, où $\alpha_n$ sont des nombres rationnels, alors la continuité de la fonction exponentielle nous donne

$$a^{xy} = \lim_{n\to\infty} a^{x\alpha_n} = \lim_{n\to\infty} (a^x)^{\alpha_n} = (a^x)^y.$$

Ceci prouve 7) pour $y > 0$.

Si $y < 0$, alors $(y = -|y|)$

$$a^{xy} = a^{-x|y|} = \frac{1}{a^{x|y|}} = \frac{1}{(a^x)^{|y|}} = (a^x)^{-|y|} = (a^x)^y.$$

Si $a \in ]0, 1[$, on pose

$$a^x = \frac{1}{(1/a)^x}.$$

Les propriétés 1), 2) et 4) restent valables. La propriété 3) devient: $a^x > a^y$ $(x < y)$.

La propriété 5): $a^x \to 0$, $x \to +\infty$.

La propriété 6): $a^x \to +\infty$, $x \to -\infty$.

d) F o n c t i o n $\log_a x$. On admettra que $a > 1$. Etant continue, strictement croissante sur $]-\infty, \infty[$ et appliquant $]-\infty, \infty[$ sur $]0, \infty[$, la fonction $y = a^x$ admet une fonction réciproque continue et strictement croissante sur $]0, \infty[$. On l'appelle *logarithme de y à base a* et on la note

$$x = \log_a y, \quad y \in ]0, \infty[.$$

De ce qui précède il s'ensuit que

$$\lim_{x\to+\infty} \log_a x = +\infty, \quad \lim_{\substack{x\to 0\\ x>0}} \log_a x = -\infty.$$

Pour $a < 1$, les raisonnements sont analogues. La fonction $a^x$ transforme aussi l'intervalle $]-\infty, \infty[$ en $]0, \infty[$, mais elle est strictement décroissante. La fonction réciproque $\log_a x$ définie sur $]0, \infty[$ sera aussi strictement décroissante et telle que

$$\lim_{x\to+\infty} \log_a x = -\infty, \quad \lim_{\substack{x\to 0\\ x>0}} \log_a x = +\infty.$$

On a (cf. § 6)

$$a^{\log_a x} = x, \quad x \in ]0, +\infty[,$$
$$\log_a a^x = x, \quad x \in ]-\infty, +\infty[, \; a \neq 1. \tag{17}$$

D'où l'on déduit, en vertu des propriétés de la fonction $a^x$, pour $x$, $y > 0$ que

$$a^{\log_a (xy)} = xy = a^{\log_a x} a^{\log_a y} = a^{\log_a x + \log_a y}$$

et

$$\log_a (xy) = \log_a x + \log_a y.$$

En remplaçant $x$ par $x/y$, on obtient

$$\log_a x - \log_a y = \log_a (x/y).$$

Par ailleurs (voir 7))

$$a^{\log_a x^y} = x^y = (a^{\log_a x})^y = a^{y \log_a x} \quad (x > 0),$$

donc

$$\log_a x^y = y \log_a x \quad (a \neq 1, \; x > 0). \tag{18}$$

Notons enfin que pour des nombres strictement positifs $a$ et $b$ non égaux à 1, on a

$$a^{\log_a b \cdot \log_b a} = (a^{\log_a b})^{\log_b a} = b^{\log_b a} = a,$$

et par suite

$$\log_a b \cdot \log_b a = 1.$$

Le logarithme du nombre $a$ à base $e$ s'appelle *logarithme naturel* ou *népérien* de $a$ et se note: $\log_e a = \ln a$.

e) Revenons à la fonction puissance

$$y = x^n, \ x \in ]0, \infty[.$$

On vient de voir que cette fonction a un sens pour les $n$ rationnels et irrationnels. On peut encore l'écrire sous la forme (cf. (17) et (18)):

$$x^n = e^{n \ln x}, \tag{19}$$

d'où il ressort qu'elle est continue en tant que composée de deux fonctions continues.

Pour $n > 0$, elle est strictement croissante et telle que

$$\lim_{\substack{x \to 0 \\ x > 0}} x^n = 0, \quad \lim_{x \to +\infty} x^n = +\infty.$$

Si l'on admet que $0^n = 0$ pour $n > 0$, alors cette fonction est continue à droite au point $x = 0$.

Pour $n < 0$, la fonction $x^n$ est continue, strictement décroissante sur $]0, \infty[$ et telle que

$$\lim_{\substack{x \to 0 \\ x > 0}} x^n = +\infty, \quad \lim_{x \to +\infty} x^n = 0.$$

De la formule (19) on déduit une propriété caractéristique de la fonction puissance

$$(xy)^n = e^{n \ln (xy)} = e^{n \ln x} e^{n \ln y} = x^n y^n \quad (x, y > 0).$$

f) F o n c t i o n $y = u(x)^{v(x)}$. Soient $u(x)$ et $v(x)$ des fonctions définies au voisinage d'un point $a$, sauf éventuellement en $a$. et supposons que $u(x) > 0$, $\lim_{x \to a} u(x) = A > 0$ et $\lim_{x \to a} v(x) = B$ ($A$ et $B$ sont des nombres finis). Alors

$$\lim_{x \to a} u(x)^{v(x)} = A^B. \tag{20}$$

En effet (voir (17), (18))

$$\lim_{x \to a} u(x)^{v(x)} = \lim_{x \to a} e^{v(x) \ln u(x)} = e^{\lim\limits_{x \to a} [v(x) \ln u(x)]} =$$

$$= e^{\lim\limits_{x \to a} v(x) \lim\limits_{x \to a} \ln u(x)} = e^{B \ln A} = A^B.$$

Dans la deuxième égalité on s'est servi de la continuité de la fonction $e^z$, dans la quatrième, de celle de la fonction ln $z$.

Si $u(x)$ et $v(x)$ sont continues en $x = a$ et $u(a) > 0$, alors $u(x) > 0$ dans un voisinage de ce point (voir théorème 4 du § 3) et $A = u(a)$, $B = v(a)$. Donc, en vertu de (20)

$$\lim_{x \to a} u(x)^{v(x)} = u(a)^{v(a)}.$$

Signalons les cas intéressants, non envisagés par l'égalité (20), où ($x \to a$, $u > 0$) $u \to +\infty$, $v \to 0$; $v \to \infty$, $u \to 1$; $v \to 0$, $u \to 0$. Le théorème de la limite de $v \ln u$ est mis en défaut dans ces cas. Si l'on ne dispose pas à priori d'une information plus exacte sur le caractère de la tendance de $u$ et de $v$ vers ces limites, on ne peut définir $\lim\limits_{x \to a} u^v$. Au voisinage de $a$ on obtient alors les formes indéterminées $\infty^0$, $1^\infty$, $0^0$.

g) Fonctions trigonométriques. Les fonctions $\sin x$, $\cos x$, $\operatorname{tg} x$ et autres sont connues du cours de trigonométrie où elles ont été définies à partir de considérations géométriques. On se basera aussi sur ces définitions.

On aurait pu donner une définition purement analytique des fonctions trigonométriques, mais on ne le fera pas.

Signalons que la fonction $y = \sin x$ est continue (voir § 3, exemple 3), strictement croissante sur $[-\pi/2, \pi/2]$ qu'elle transforme en $[-1, 1]$. Elle admet donc une fonction réciproque continue

$$x = \operatorname{Arcsin} y, \quad y \in [-1, 1].$$

Cependant, sur l'axe des $x$ tout entier $]-\infty, \infty[$ la réciproque de la fonction $y = \sin x$ est la correspondance infinivalente $\arcsin y$ dont les valeurs se calculent par la formule

$$x = \arcsin y = (-1)^k \operatorname{Arcsin} y + k\pi \quad (k = 0, \pm 1, \pm 2, \ldots), \tag{21}$$

c'est-à-dire qu'à chaque $y \in [-1, 1]$ est associé d'ensemble $e_y$ des valeurs de $x$ définies par la formule (21).

De façon identique, les réciproques des fonctions

$$y = \cos x, \quad x \in [0, \pi],$$
$$y = \operatorname{tg} x, \quad x \in ]-\pi/2, \pi/2[,$$

sont les fonctions

$$x = \operatorname{Arccos} y, \quad y \in [-1, 1],$$
$$x = \operatorname{Arctg} y, \quad y \in ]-\infty, \infty[.$$

Si l'on considère les fonctions $y = \cos x$ et $y = \operatorname{tg} x$ sur l'axe des $x$ tout entier, leurs réciproques respectives seront les correspondances

$$x = \arccos y = \pm \operatorname{Arccos} y + 2k\pi,$$
$$x = \operatorname{arctg} y = \operatorname{Arctg} y + k\pi$$
$$(k = 0, \pm 1, \pm 2, \ldots).$$

f) Fonctions hyperboliques. Les fonctions

$$\text{sh}\, x = \frac{e^x - e^{-x}}{2}, \quad \text{ch}\, x = \frac{e^x + e^{-x}}{2}, \quad \text{th}\, x = \frac{\text{sh}\, x}{\text{ch}\, x}, \quad \text{coth}\, x = \frac{\text{ch}\, x}{\text{sh}\, x}$$

sont dites respectivement *sinus hyperbolique*, *cosinus hyperbolique*, *tangente hyperbolique* et *cotangente hyperbolique*. Leurs graphes sont représentés sur les figures 32 et 33. Les fonctions sh $x$, ch $x$ et th $x$

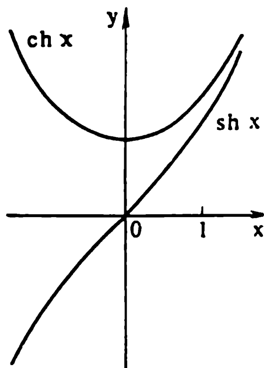



Fig. 32

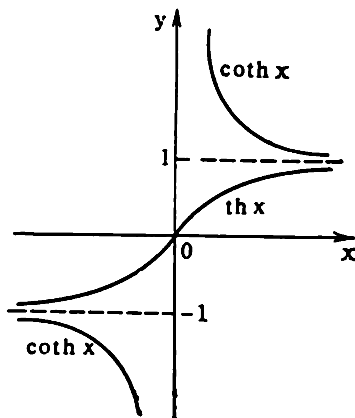



Fig. 33

sont définies sur $]-\infty, \infty[$, la fonction coth $x$ est définie sur le même intervalle privé du point 0.

Il est immédiat de vérifier que ces fonctions sont justiciables de formules similaires (mais pas toujours identiques) à celles de la trigonométrie. Par exemple

$$\text{sh}\,(x + y) = \text{sh}\, x\, \text{ch}\, y + \text{sh}\, y\, \text{ch}\, x,$$

$$\text{ch}\,(x + y) = \text{ch}\, x\, \text{ch}\, y + \text{sh}\, x\, \text{sh}\, y.$$

En faisant $y = -x$ dans la dernière relation, on obtient

$$\text{ch}^2 x - \text{sh}^2 x = 1.$$

Les fonctions hyperboliques envisagées sont toutes continues dans leur domaine de définition. Les fonctions réciproques de sh $x$ et th $x$ s'appellent *argument sinus hyperbolique* et *argument tangente hyperbolique* et se notent Arg sh $y$ et Arg th $y$. Pour $x \geqslant 0$, la réciproque de ch $x$ est aussi une fonction.

## § 9. Limites remarquables

THÉORÈME 1.

$$\lim_{x\to 0}\frac{\sin x}{x}=1.$$

DÉMONSTRATION. Etant continue, la fonction $y = \sin x$ tend vers 0 avec $x$. Donc on obtient un cas d'indétermination de la forme $\left(\frac{0}{0}\right)$. Levons cette indétermination. Par définition des fonctions trigonométriques et par des considérations géométriques on a (fig. 34)

$$0 < \sin x < x < \operatorname{tg} x$$

pour $0 < x < \pi/2$ ($MN = \sin x$, $AM \perp OM$, $AM = \operatorname{tg} x$, $OM = 1$). En divisant par $\sin x > 0$, on obtient

$$1<\frac{x}{\sin x}<\frac{1}{\cos x} \quad \text{ou} \quad 1>\frac{\sin x}{x}>\cos x. \tag{1}$$

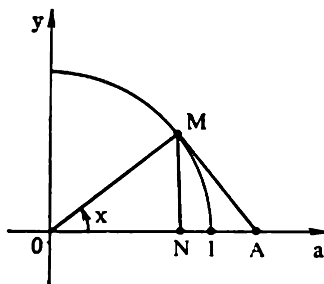



Fig. 34

Les inégalités (1) sont valables pour $x \in\, ]-\pi/2, 0[$, puisque les fonctions $\cos x$ et $\frac{\sin x}{x}$ sont paires. Par ailleurs, la fonction $\cos x$ étant continue, on a

$$\lim_{x\to 0}\cos x = \cos 0 = 1,$$

donc, en passant à la limite dans (1) et compte tenu du théorème 4 du § 2, il vient que

$$\lim_{x\to 0}\frac{\sin x}{x}=1.$$

EXEMPLE 1. $\lim\limits_{x\to 0}\frac{1-\cos x}{x^2}=\lim\limits_{x\to 0}\frac{2\sin^2\frac{x}{2}}{x^2}=\frac{1}{2}\lim\limits_{x\to 0}\left(\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2=$

$$=\frac{1}{2}\left(\lim_{x\to 0}\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2=\frac{1}{2}\cdot 1=\frac{1}{2}.$$

THÉORÈME 2.

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e. \tag{2}$$

DÉMONSTRATION. Par définition de la limite d'une fonction, on doit prouver que

$$\left(1+\frac{1}{x_n}\right)^{x_n} \to e, \quad \forall x_n \to \infty. \tag{3}$$

Si $x_n$ est un entier naturel, ceci a déjà été prouvé (voir § 6, chap. 2). Assurons-nous maintenant que la relation (2) est vraie lorsque $x_n \to +\infty$ et $x_n \to -\infty$ par valeurs non nécessairement entières (voir remarque à la fin du § 2).

Soit $x_n$ une variable tendant vers $+\infty$ et posons $[x_n] = k_n$. On a alors $k_n \leqslant x_n < k_n + 1 \leqslant x_n + 1 < k_n + 2$ et

$$\left(1+\frac{1}{k_n+1}\right)^{k_n+1} < \left(1+\frac{1}{x_n}\right)^{x_n+1} < \left(1+\frac{1}{k_n}\right)^{k_n+2} < e\left(1+\frac{1}{k_n}\right)^2.$$

Lorsque $x_n \to +\infty$, la partie entière $[x_n] = k_n \to +\infty$, donc le premier et le dernier membre tendent vers $e$. Par suite

$$\left(1+\frac{1}{x_n}\right)^{x_n+1} \to e,$$

et comme $1+\frac{1}{x_n} \to 1$, on a prouvé (3) pour $x_n \to +\infty$.

Si maintenant $x_n \to -\infty$, alors $x'_n = -x_n \to +\infty$ et

$$\lim_{x_n \to -\infty} \left(1+\frac{1}{x_n}\right)^{x_n} = \lim_{x'_n \to +\infty} \left(1-\frac{1}{x'_n}\right)^{-x'_n} = \lim_{x'_n \to +\infty} \left(\frac{x'_n}{x'_n-1}\right)^{x'_n} =$$

$$= \lim_{x'_n \to +\infty} \left[\left(1+\frac{1}{x'_n-1}\right)^{x'_n-1} \left(1+\frac{1}{x'_n-1}\right)\right] = e,$$

ce qui prouve (2).

EXEMPLE 2.

$$\lim_{u \to 0} (1+u)^{1/u} = e.$$

Ce résultat se déduit à partir de (2) par la substitution $1/x = u$.

EXEMPLE 3.

$$\lim_{x \to \infty} \left(1+\frac{\alpha}{x}\right)^x = e^\alpha, \quad \lim_{u \to 0} (1+\alpha u)^{1/u} = e^\alpha, \quad \forall \alpha.$$

Pour $\alpha = 0$, on a $\lim\limits_{x \to \infty} 1^x = 1 = e^0$, car par définition $1^x = 1$.

Supposons maintenant que $\alpha \neq 0$. Si $x \to \infty$, alors $\frac{x}{\alpha} \to \infty$ et

$$\left(1+\frac{\alpha}{x}\right)^x = \left[\left(1+\frac{\alpha}{x}\right)^{x/\alpha}\right]^\alpha = u^\alpha \underset{u \to e}{\longrightarrow} e^\alpha,$$

puisque la fonction puissance $u^\alpha$ est continue au point $u = e$ (cf. § 8, e)).

EXEMPLE 4.

$$\lim_{x \to 0} \frac{\ln(1+x)}{x} = 1, \quad \lim_{x \to 0} \frac{\log_a(1+x)}{x} = \log_a e = \frac{\phantom{1}}{\ln a}$$

La fonction $\ln u$ étant continue sur l'intervalle $]0, \infty[$, on a (voir exemple 2)

$$\lim_{x \to 0} \frac{\ln(1+x)}{x} = \lim_{x \to 0} \ln(1+x)^{1/x} = \ln \lim_{x \to 0} (1+x)^{1/x} = \ln e = 1.$$

EXEMPLE 5.

$$\lim_{x \to 0} \frac{a^x - 1}{x} = \ln a \quad (a > 0), \quad \lim_{x \to 0} \frac{e^x - 1}{x} = 1.$$

Posons en effet $a^x - 1 = u$. La fonction puissance étant continue, $u \to 0$ avec $x$. D'autre part, $x \ln a = \ln(1+u)$, donc (voir exemple 4)

$$\lim_{x \to 0} \frac{a^x - 1}{x} = \lim_{u \to 0} \frac{u}{\ln(1+u)} \ln a = \ln a \cdot \lim_{u \to 0} \frac{u}{\ln(1+u)} = \ln a.$$

## § 10. Ordre d'une variable. Equivalence

Considérons deux fonctions $\varphi(x)$ et $\psi(x)$ définies dans un voisinage $V(a)$ d'un point $a$, sauf éventuellement en $a$, qui peut être fini ou infini. On admettra de plus que $\psi(x) \neq 0$ sur $V(a)$. Si

$$\lim_{x \to a} \frac{\varphi(x)}{\psi(x)} = 0, \tag{1}$$

on note ce fait par:

$$\varphi(x) = o(\psi(x)), \quad x \to a, \tag{1'}$$

et on dit que $\varphi(x)$ est égal à petit $o$ de $\psi(x)$ pour $x \to a$. On dit encore que $\varphi(x)$ *est négligeable devant* $\psi(x)$ ou que $\psi(x)$ *est prépondérante sur* $\varphi(x)$.

Par exemple:

$$\begin{array}{lll} x^2 = o(x), & x \to 0; & \\ x^n = o(x^m), & x \to 0 & \text{pour } m < n; \\ x^n = o(x^m), & x \to \infty & \text{pour } m > n; \\ (x-a)^4 = o((x-a)^3), & x \to a; & \\ 1 - \cos x = o(x), & x \to 0, & \text{car } \lim\limits_{x \to 0} \dfrac{1 - \cos x}{x} = 0. \end{array}$$

L'expression $o(1)$, $x \to a$, représente un infiniment petit pour $x \to a$, c'est-à-dire une fonction $\varphi(x)$ qui tend vers 0 pour $x \to a$. Par exemple, $\varphi(x) = \frac{1}{\ln x} = o(1)$, $x \to +\infty$.

La relation (1) exprime visiblement le fait que la fonction $\varphi(x)$ peut s'écrire sous la forme: $\varphi(x) = \varepsilon(x)\psi(x)$, où $\varepsilon(x) \to 0$ pour $x \to a$.

Si les fonctions $\varphi$ et $\psi$ de (1) (ou de (1')) sont des infiniment petits pour $x \to a$, on dit, suivant une terminologie plus ancienne, que $\varphi(x)$ *pour* $x \to a$ *est un infiniment petit d'ordre supérieur à* $\psi(x)$.

Si $\varphi$ et $\psi$ sont des infiniment grands pour $x \to a$, on dit alors que $\varphi(x)$ *est un infiniment grand d'ordre inférieur à* $\psi(x)$ ou encore que $\psi(x)$ *est un infiniment grand d'ordre supérieur à* $\varphi(x)$.

On écrira encore

$$\varphi(x) \approx \psi(x), \quad x \to a, \tag{2}$$

et on dira que les fonctions $\varphi(x)$ et $\psi(x)$ sont *équivalentes* (*asymptotiquement égales*) *pour* $x \to a$ si

$$\lim_{x\to a} \frac{\varphi(x)}{\psi(x)} = 1. \tag{2'}$$

Par exemple

$$\sin x \approx x, \quad x \to 0, \tag{3}$$

ou encore (voir exemples 1, 4, 5 du § 9)

$$1 - \cos x \approx x^2/2, \qquad x \to 0, \tag{4}$$

$$\ln(1 + x) \approx x, \qquad x \to 0, \tag{5}$$

$$e^x - 1 \approx x, \qquad x \to 0, \tag{6}$$

$$a^x - 1 \approx x \ln a, \qquad x \to 0. \tag{7}$$

Signalons que

$$\lim_{x\to a} f(x) = A \neq 0$$

équivaut à

$$\lim_{x\to a} \frac{f(x)}{A} = 1,$$

ce que nous avons convenu de noter aussi :

$$f(x) \approx A, \quad x \to a \quad (A \neq 0). \tag{8}$$

La terminologie introduite nous permettra d'alléger les calculs et l'écriture des formules. Il importe pour cela d'assimiler quelques propriétés simples des fonctions équivalentes, propriétés qui sont exprimées dans les théorèmes qui suivent.

THEOREME 1. *Si*

$$\varphi(x) \approx \psi(x), \qquad x \to a, \tag{9}$$

*alors*

$$\psi(x) \approx \varphi(x), \qquad x \to a. \tag{10}$$

Démonstration. Si $\psi(x) \neq 0$ sur $V(a)$ et la relation (9) a lieu, alors il est évident que $\varphi(x) \neq 0$ éventuellement dans un voisinage moindre. Mais alors

$$\lim_{x \to a} \frac{\psi(x)}{\varphi(x)} = \lim_{x \to a} \frac{1}{\frac{\varphi(x)}{\psi(x)}} = \frac{1}{1} = 1.$$

Théorème 2. *Pour que la relation* (9) *ait lieu, il est nécessaire et suffisant que*

$$\varphi(x) = \psi(x) + o(\psi(x)), \quad x \to a. \tag{11}$$

L'égalité (11) doit être lue de la manière suivante : le terme ajouté à $\psi(x)$ pour obtenir $\varphi(x)$ est tel que son quotient par $\psi(x)$ tend vers 0 pour $x \to a$.

Démonstration. Supposons que la relation (9) est vérifiée. Alors

$$\frac{\varphi(x)}{\psi(x)} = 1 + \varepsilon(x), \quad \text{où} \quad \varepsilon(x) \to 0 \quad \text{pour } x \to a.$$

Donc,

$$\varphi(x) = \psi(x) + \varepsilon(x)\psi(x) = \psi(x) + o(\psi(x)), \ x \to a,$$

ce qui prouve (11).

Réciproquement, si (11) est réalisée, alors

$$\varphi(x) = \psi(x) + o(\psi(x)) = \psi(x) + \varepsilon(x)\psi(x),$$

où $\varepsilon(x) \to 0$ pour $x \to a$. Donc,

$$\frac{\varphi(x)}{\psi(x)} = 1 + \varepsilon(x) \xrightarrow[x \to a]{} 1,$$

et l'on obtient (9).

Théorème 3. *Si*

$$\psi(x) \approx \psi_1(x), \quad x \to a,$$

*alors*

$$\lim_{x \to a} [\varphi(x)\psi(x)] = \lim_{x \to a} [\varphi(x)\psi_1(x)], \tag{12}$$

$$\lim_{x \to a} \frac{\varphi(x)}{\psi(x)} = \lim_{x \to a} \frac{\varphi(x)}{\psi_1(x)}. \tag{13}$$

*Ces égalités doivent être comprises au sens que l'existence de l'une des limites entraîne celle de l'autre et ces limites sont égales.*

D'où il s'ensuit que si l'une de ces limites n'existe pas, il en est de même de l'autre.

Démonstration. Prouvons la première relation. Supposons que la limite à droite existe dans (12). Alors

$$\lim_{x\to a} [\varphi(x)\psi(x)] = \lim_{x\to a}\left[\varphi(x)\psi_1(x)\frac{\psi(x)}{\psi_1(x)}\right] =$$

$$= \lim_{x\to a} [\varphi(x)\psi_1(x)] \lim_{x\to a}\frac{\psi(x)}{\psi_1(x)} = \lim_{x\to a} [\varphi(x)\psi_1(x)]\cdot 1 =$$

$$= \lim_{x\to a} [\varphi(x)\psi_1(x)].$$

Exemple 1. $\operatorname{tg} x \approx x$, $x \to 0$, car

$$\lim_{x\to 0}\frac{\operatorname{tg} x}{x} = \lim_{x\to 0}\left(\frac{\sin x}{x}\cdot\frac{1}{\cos x}\right) = 1.$$

Exemple 2.

$$\lim_{x\to 0}\frac{\operatorname{tg}^3 x}{x^3+x} = \lim_{x\to 0}\frac{x^3}{x^3+x} = \lim_{x\to 0}\frac{x^2}{x^2+1} = \frac{0}{1} = 0.$$

Définition. *Etant donné une fonction $\varphi(x)$, si l'on peut choisir des nombres $A$ et $m$, où $A \neq 0$, tels que*

$$\varphi(x) \approx A(x-a)^m, \quad x \to a,$$

*on dit que la fonction $A(x-a)^m$ est le terme principal de la fonction $\varphi(x)$ au voisinage de $a$.*

Les seconds membres des relations (3) à (7) sont de toute évidence des termes principaux pour $x \to 0$.

On dira qu'une fonction $f$ est *du même ordre* qu'une fonction $\varphi$ sur un ensemble $E$, ou encore que $f$ est *dominée par* $\varphi$, si

$$|f(x)| \leqslant C\,|\varphi(x)|, \quad \forall x \in E,$$

$C$ étant une constante strictement positive ne dépendant pas de $x$, et l'on notera

$$f(x) = O(\varphi(x)). \tag{14}$$

Le symbole $O(\varphi(x))$ se lit naturellement: grand $O$ de $\varphi$.

En particulier, la relation

$$f(x) = O(1) \text{ sur } E$$

signifie que $f$ est bornée sur $E$.

Exemples:

1) $\sin x = O(1)$, $\sin x = O(x)$, sur $]-\infty, \infty[$;

2) $x = O(x^2)$ sur $[1, \infty[$;

3) $x^2 = O(x)$ sur $[0, 1]$.

CHAPITRE 4

# CALCUL DIFFÉRENTIEL POUR FONCTIONS D'UNE VARIABLE

## § 1. Dérivée

La notion de dérivée est primordiale en analyse mathématique.

*On appelle dérivée d'une fonction f en un point x la limite du rapport de son accroissement $\Delta y$ en ce point à l'accroissement correspondant $\Delta x$ lorsque ce dernier tend vers* 0.

La dérivée se note:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \qquad (1)$$

On se sert aussi des notations: $y'$, $\frac{df(x)}{dx}$, $\frac{dy}{dx}$. La commodité de ces notations apparaîtra plus loin.

Pour $x$ fixe, la quantité $\frac{\Delta y}{\Delta x}$ est une fonction de $\Delta x$:

$$\psi(\Delta x) = \frac{\Delta y}{\Delta x} \quad (\Delta x \neq 0).$$

Une condition nécessaire d'existence de la dérivée de $f$ en un point $x$ est que la fonction $f$ soit définie dans un voisinage de $x$, y compris en $x$. Alors la fonction $\psi(\Delta x)$ est définie pour les $\Delta x$ assez petits non nuls, c'est-à-dire pour les $\Delta x$ vérifiant les inégalités $0 < |\Delta x| < \delta$, où $\delta$ est assez petit.

La limite (1) n'existe pas pour toute fonction $f$ définie au voisinage d'un point $x$. En général, quand on dit qu'une fonction $f$ admet en un point $x$ une dérivée $f'(x)$, on sous-entend que cette dérivée est finie, c'est-à-dire que la limite (1) est finie. Cependant, il est fort possible que la limite (1) soit égale à $\pm\infty$. Dans ces cas on dit que la fonction $f$ admet une *dérivée infinie* en $x$.

Si dans la formule (1) on admet que $\Delta x$ tend vers 0 par valeurs positives, alors la limite correspondante (si elle existe) s'appelle *dérivée à droite de f en x* et se note $f'_d(x)$.

De façon analogue, lorsque $\Delta x$ tend vers 0 par valeurs négatives, la limite (1) s'appelle *dérivée à gauche de la fonction f en x* et se note $f'_g(x)$.

Pour calculer $f'_d(x)$ (resp. $f'_g(x)$) il est nécessaire que la fonction $f$ soit donnée en $x$ et dans un voisinage à droite (resp. à gauche) de $x$.

Un cas typique est celui où $f$ est définie sur un intervalle $[a, b]$ et admet une dérivée en tous les points intérieurs de $[a, b]$, c'est-à-dire en tous les points de $]a, b[$, une dérivée à droite en $a$ et une dérivée à gauche en $b$. Dans ce cas, on dit que la *fonction $f$ admet une dérivée sur l'intervalle* $[a, b]$ sans spécifier qu'en $a$ c'est une dérivée à droite et en $b$, une dérivée à gauche.

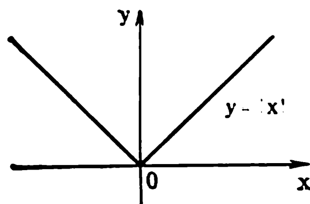


Fig. 35

Il est aisé de voir que si une fonction $f$ admet une dérivée à droite et une dérivée à gauche en $x$ égales, alors elle admet une dérivée en $x$:

$$f'_d(x) = f'_g(x) = f'(x).$$

Si une fonction $f$ admet une dérivée à droite et une dérivée à gauche en $x$ telles que $f'_d(x) \neq f'_g(x)$, alors elle n'admet pas de dérivée en $x$.

Exemple. Soit la fonction $y = |x|$ (fig. 35). On a

$$\frac{\Delta y}{\Delta x} = \frac{|x+\Delta x| - |x|}{\Delta x}.$$

Si $x > 0$, alors $x + \Delta x > 0$ pour $\Delta x$ assez petits et

$$\frac{\Delta y}{\Delta x} = \frac{x+\Delta x - x}{\Delta x} = \frac{\Delta x}{\Delta x} = 1.$$

Si $x < 0$, alors $x + \Delta x < 0$ pour $\Delta x$ assez petits et

$$\frac{\Delta y}{\Delta x} = \frac{-(x+\Delta x)-(-x)}{\Delta x} = -\frac{\Delta x}{\Delta x} = -1.$$

Donc

$$|x|' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \begin{cases} 1 & \text{pour } x > 0, \\ -1 & \text{pour } x < 0. \end{cases}$$

Supposons maintenant que $x = 0$. Alors

$$\frac{\Delta y}{\Delta x} = \frac{|\Delta x|}{\Delta x} = \operatorname{sgn} \Delta x \frac{\Delta x}{\Delta x} = \operatorname{sgn} \Delta x = \begin{cases} 1 & \text{pour } \Delta x > 0, \\ -1 & \text{pour } \Delta x < 0. \end{cases}$$

Donc

$$\lim_{\substack{\Delta x \to 0 \\ \Delta x > 0}} \frac{\Delta y}{\Delta x} = 1, \quad \lim_{\substack{\Delta x \to 0 \\ \Delta x < 0}} \frac{\Delta y}{\Delta x} = -1.$$

Ainsi, la fonction $|x|$ admet en $x = 0$ une dérivée à droite égale à 1 et une dérivée à gauche égale à $-1$, donc elle n'admet pas de *dérivée* en $x = 0$.

Nous savons (voir chap. 3, § 3, exemple 8) que la fonction $|x|$ est continue pour toutes les valeurs de $x$, y compris pour $x = 0$. Donc, elle peut servir d'exemple de fonction partout continue n'admettant pas de dérivée en un point. Il existe d'autres exemples de fonctions continues sur l'axe des réels tout entier et non dérivables sur cet axe. C'est le cas notamment de la fonction de Weierstrass

$$\mu(x) = \sum_{n=0}^{\infty} 2^{-n} \sin 2^n \pi x.$$

Passons sur les détails de cette question.

D'autre part, *toute fonction admettant une dérivée (finie!) en un point $x$ est continue en ce point.*

En effet, supposons que la limite (1) existe en un point $x$ et est finie. Ce fait se note:

$$\frac{\Delta y}{\Delta x} = f'(x) + \varepsilon(\Delta x), \tag{2}$$

où $\varepsilon(\Delta x) \to 0$ pour $\Delta x \to 0$, c'est-à-dire que $\varepsilon(\Delta x)$ est un infiniment petit pour $\Delta x \to 0$. De (2) il s'ensuit que:

$$\Delta y = f'(x)\,\Delta x + \Delta x \cdot \varepsilon(\Delta x).$$

En passant à la limite lorsque $\Delta x \to 0$, on obtient

$$\lim_{\Delta x \to 0} \Delta y = 0,$$

ce qui montre que $f$ est continue en $x$.

Voici quelques applications importantes de la dérivée.

V i t e s s e   i n s t a n t a n é e. Supposons que la fonction $s = f(t)$ décrit le mouvement d'un point sur une droite assimilée à l'axe des $s$; $s$ représente donc l'abscisse du point mobile à l'instant $t$. Le chemin parcouru entre $t$ et $t + \Delta t$ est

$$\Delta s = f(t + \Delta t) - f(t).$$

La vitesse moyenne est alors

$$v_m = \frac{\Delta s}{\Delta t}.$$

La *vitesse instantanée* à l'instant $t$ se définit naturellement comme la limite

$$v = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = f'(t).$$

Intensité du courant. Soit $Q = f(t)$ la quantité de courant traversant un fil pendant une durée de temps $t$. Alors

$$\frac{\Delta Q}{\Delta t} = \frac{f(t+\Delta t) - f(t)}{\Delta t}$$

est l'intensité moyenne du courant pendant l'intervalle $[t, t + \Delta t]$. Quant à la limite

$$\lim_{\Delta t \to 0} \frac{\Delta Q}{\Delta t} = Q' = I,$$

elle représente l'*intensité du courant à l'instant* $t$.

Densité de répartition d'une masse. Supposons qu'une masse est répartie de façon non uniforme sur un intervalle $[a, b]$ de l'axe $Ox$ (fig. 36) et de telle sorte que la quantité de masse affectée à $[a, x]$ soit égale à

$$M = F(x) \quad (a \leqslant x \leqslant b).$$

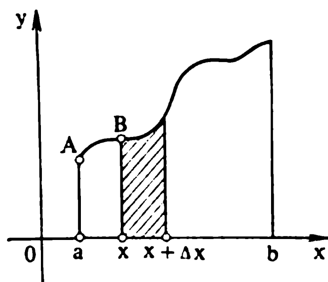


Fig. 36

Cette quantité est proportionnelle à l'aire de la figure $aABx$. Donc, $M$ est une fonction de $x$ ($M = F(x)$). La quantité de masse affectée à l'intervalle $[x, x + \Delta x]$ est visiblement égale à

$$\Delta F = F(x + \Delta x) - F(x).$$

La densité moyenne sur cet intervalle est égale à $\frac{\Delta F}{\Delta x}$, la limite

$$\lim_{\Delta x \to 0} \frac{\Delta F}{\Delta x} = F'(x) = \mu$$

est la *densité réelle de répartition de la masse en* $x$.

## § 2. Signification géométrique de la dérivée

Soit donnée une fonction $y = f(x)$ continue sur un intervalle $]a, b[$. Son graphe s'appelle *courbe continue*. Désignons-le par $\Gamma$. Repérons un point $A = (x, f(x))$ de $\Gamma$ (fig. 37 et 38) et définissons la tangente à $\Gamma$ en $A$. Considérons à cet effet un autre point $B = (x + \Delta x, f(x + \Delta x))$ de $\Gamma$, où $\Delta x \neq 0$ (la figure 37 correspond au cas où $\Delta x > 0$, la figure 38, au cas où $\Delta x < 0$). Appelons *sécante* et désignons-la par $S$ la droite passant par $A$ et $B$ et orientée dans le sens des $x$ croissants. Soit $\beta$ l'angle de $S$ avec $Ox$. On admet que $-\pi/2 < \beta < \pi/2$. Sur la figure 37, on a $\Delta x = AC$, $\Delta y = CB$, sur la figure 38, $\Delta x = -AC$, $\Delta y = -CB$. Dans les deux cas $\Delta y/\Delta x = \operatorname{tg} \beta$.

Si $\Delta x \to 0$, alors $\Delta y \to 0$ et le point $B$ tend vers $A$ le long de $\Gamma$. Ceci étant, si l'angle $\beta$ tend vers une valeur $\alpha$ distincte de $\pi/2$ et de $-\pi/2$, alors existe la limite

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\beta \to \alpha} \operatorname{tg} \beta = \operatorname{tg} \alpha, \tag{1}$$

dans laquelle on reconnaît la dérivée de $f$ en $x$:

$$f'(x) = \operatorname{tg} \alpha. \tag{2}$$

Inversement, si existe la dérivée (finie) $f'(x)$, alors $\beta \to \alpha = \operatorname{arctg} f'(x)$.

Lorsque $\beta$ tend vers $\alpha$, la sécante $S$ tend vers la droite orientée $T$ qui passe par $A$ et fait un angle $\alpha$ avec l'axe $Ox$.

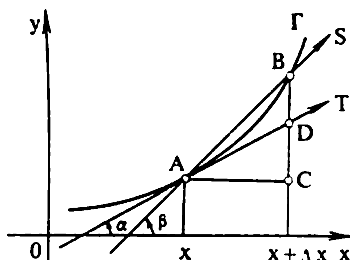



Fig. 37

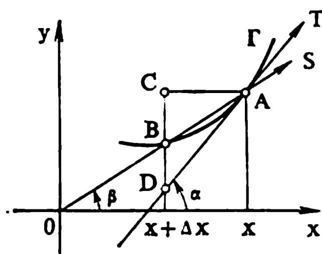



Fig. 38

La droite orientée $T$ s'appelle *tangente à la courbe* $\Gamma$ *au point* $A$.

DÉFINITION. *On appelle tangente à une courbe* $\Gamma$ $(y = f(x))$ *en un point* $A = (x, f(x))$ *la droite orientée* $T$ *vers laquelle tend la sécante* $S$ *(orientée dans le sens des* $x$ *croissants) qui passe par* $A$ *et par un point* $B = (x + \Delta x,\ f(x + \Delta x)) \in \Gamma$, *lorsque* $\Delta x \to 0$.

On vient de prouver que *si une fonction continue* $y = f(x)$ *admet une dérivée finie* $f'(x)$ *en un point* $x$, *alors sa courbe représentative présente en* $x$ *une tangente de coefficient angulaire* $\operatorname{tg} \alpha = f'(x)$ $(-\pi/2 < \alpha < \pi/2)$.

Réciproquement, l'existence de la limite

$$\lim \beta = \alpha,\ \alpha \in \left]-\pi/2,\ \pi/2\right[$$

entraîne celle de la dérivée finie $f'(x)$ et les égalités (1) et (2).

Il n'est pas exclu que $f$ admette en $x$ une dérivée à droite et une dérivée à gauche distinctes:

$$f'_g(x) \neq f'_d(x).$$

On dit alors que le point $A$ est *anguleux*. Dans ce cas la courbe $\Gamma$ ne présente pas de tangente en $A$. Par contre, elle admet une demi-

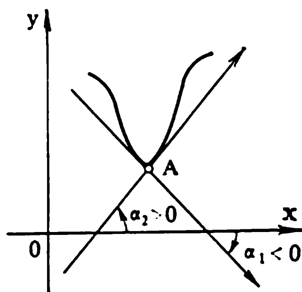



Fig. 39

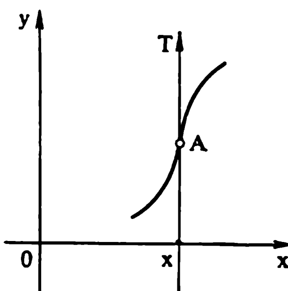



Fig. 40

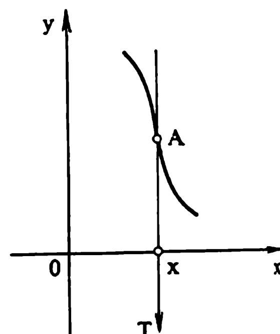



Fig. 41

tangente à droite et une autre à gauche avec des coefficients angulaires différents (fig. 39) :

$$\operatorname{tg}\alpha_1 = \lim_{\substack{\Delta x\to 0\\ \Delta x<0}} \frac{\Delta y}{\Delta x} = f'_g(x), \quad \operatorname{tg}\alpha_2 = \lim_{\substack{\Delta x\to 0\\ \Delta x>0}} \frac{\Delta y}{\Delta x} = f'_d(x).$$

Supposons maintenant que la dérivée de $f$ est infinie en un point $x$:

$$f'(x) = \lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x} = \infty.$$

Distinguons quatre cas importants :

1) $f'(x) = \lim\limits_{\Delta x\to 0} \frac{\Delta y}{\Delta x} = +\infty, \quad \beta \to \frac{\pi}{2}$ (fig. 40).

2) $f'(x) = \lim\limits_{\Delta x\to 0} \frac{\Delta y}{\Delta x} = -\infty, \quad \beta \to -\frac{\pi}{2}$ (fig. 41).

3) $f'_g(x) = \lim\limits_{\substack{\Delta x\to 0\\ \Delta x<0}} \frac{\Delta y}{\Delta x} = -\infty, \quad \beta \to -\frac{\pi}{2}.$

$f'_d(x) = \lim\limits_{\substack{\Delta x\to 0\\ \Delta x>0}} \frac{\Delta y}{\Delta x} = +\infty, \quad \beta \to \frac{\pi}{2}.$

La demi-tangente à gauche est perpendiculaire à l'axe $Ox$ et dirigée vers le bas. La demi-tangente à droite est perpendiculaire

à l'axe $Ox$ et dirigée vers le haut (fig. 42).

4) $$f'_g(x) = \lim_{\substack{\Delta x \to 0 \\ \Delta x < 0}} \frac{\Delta y}{\Delta x} = +\infty, \quad \beta \to \frac{\pi}{2},$$

$$f'_d(x) = \lim_{\substack{\Delta x \to 0 \\ \Delta x > 0}} \frac{\Delta y}{\Delta x} = -\infty, \quad \beta \to -\frac{\pi}{2}.$$

Les demi-tangentes à gauche et à droite sont parallèles à l'axe $Oy$, la première est dirigée vers le haut, la seconde vers le bas (fig. 43).

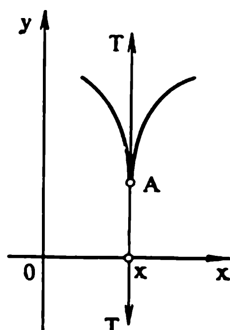



Fig. 42

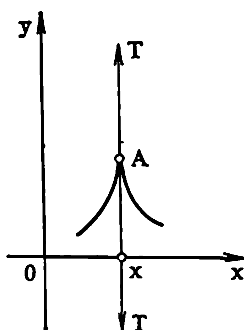



Fig. 43

Remarque. D'ordinaire on définit la tangente à une courbe de la manière suivante : on appelle tangente $T$ à une courbe $\Gamma$ en un point $A$ la droite vers laquelle tend une sécante $S$ passant par $A$ et par un point $B \in \Gamma$, lorsque ce dernier tend vers $A$ le long de $\Gamma$.

Cette définition n'implique pas que $S$ et $T$ soient des droites orientées. Elle est correcte si la tangente n'est pas parallèle à l'axe $Oy$. Si on applique cette définition au cas 4) par exemple (fig. 43, où $A$ est un point anguleux), on obtient que la courbe considérée admet une tangente unique en $A$.

La définition citée nous donne au point $A$ deux demi-tangentes de sens contraires.

Du cours de géométrie analytique on sait que l'équation d'une droite (dans le plan) passant par un point $(x_0, y_0)$ et faisant un angle $\alpha$ $(-\pi/2 < \alpha < \pi/2)$ avec l'axe $Ox$ est $y - y_0 = m(x - x_0)$, $m = \operatorname{tg} \alpha$ (voir chap. 10, § 8). Donc, l'équation de la tangente à la courbe $y = f(x)$ en $(x_0, y_0)$ s'écrit

$$y - y_0 = y'_0(x - x_0), \tag{3}$$

où $y_0 = f(x_0)$, $y'_0 = f'(x_0)$.

On appelle *normale à une courbe* $\Gamma$ *en un point* $A$ la droite perpendiculaire en $A$ à la tangente à $\Gamma$ en $A$. Son équation est visiblement de la forme

$$y-y_0=-\frac{1}{y_0'}(x-x_0). \tag{4}$$

EXEMPLE 1 [1]). Trouver l'équation de la tangente à la courbe

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad (x\in[-a,\ a]) \tag{5}$$

en un point $(x_0,\ y_0)$.

La courbe (5) est une *ellipse*. Les variables $x$ et $y$ figurant par leurs carrés dans (5), l'ellipse est symétrique par rapport aux axes de coordonnées. Pour former l'équation de la tangente, on admettra que $-a\leqslant x\leqslant a$, $0\leqslant y\leqslant b$. De (5) il vient

$$y=\frac{b}{a}\sqrt{a^2-x^2}. \tag{5'}$$

D'où

$$y'=\frac{-bx}{a\sqrt{a^2-x^2}}.$$

Calculons la fonction $y$ et la dérivée $y'$ en $x_0$:

$$y_0=y(x_0)=\frac{b}{a}\sqrt{a^2-x_0^2},\quad y'(x_0)=\frac{-bx_0}{a\sqrt{a^2-x_0^2}},$$

$$y_0y'(x_0)=-\frac{b^2}{a^2}x_0. \tag{6}$$

L'équation de la tangente à l'ellipse en $(x_0,\ y_0)$ est:

$$Y-y_0=y'(x_0)(X-x_0). \tag{7}$$

En multipliant (7) par $y_0/b^2$, on obtient en vertu de (6)

$$\frac{Yy_0}{b^2}-\frac{y_0^2}{b^2}=y'(x_0)\left(\frac{Xy_0}{b^2}-\frac{x_0y_0}{b^2}\right),\quad \frac{Yy_0}{b^2}-\frac{y_0^2}{b^2}=-\frac{Xx_0}{a^2}+\frac{x_0^2}{a^2}.$$

Comme $\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}=1$, l'équation de la tangente s'écrit:

$$\frac{Xx_0}{a^2}+\frac{Yy_0}{b^2}=1. \tag{8}$$

Donc, pour former l'équation de la tangente à l'ellipse en un point $(x_0,\ y_0)$ il suffit de substituer dans (5) $Xx_0$ à $x^2$ et $Yy_0$ à $y^2$.

On raisonne de même pour les valeurs négatives de $y$ $(-b\leqslant y<$ $<0)$ et l'on établit finalement que (8) est l'équation de la tangente

[1]) Les exemples 1, 2 et 3 peuvent être traités au § 8 après l'assimilation des techniques de dérivation.

en tout point $(x_0, y_0)$ de l'ellipse. De (8) on voit que la tangente à l'ellipse en $(x_0, y_0)$ coupe l'axe $Ox$ au point d'abscisse $a^2/x_0$, point situé à droite de l'ellipse pour $x_0 > 0$ et à gauche pour $x_0 < 0$ (fig. 44).

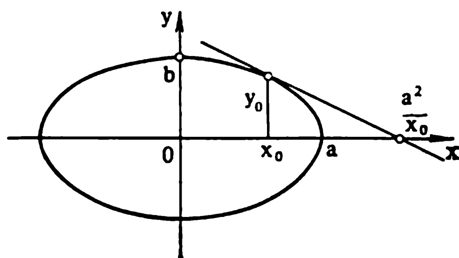



Fig. 44

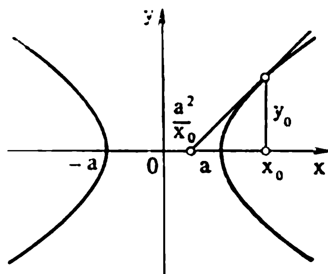



Fig. 45

Exemple 2. Trouver l'équation de la tangente à la courbe

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (|x| \geqslant a) \tag{9}$$

en un point $(x_0, y_0)$.

La courbe (9) est une *hyperbole*. Elle est aussi symétrique par rapport aux axes de coordonnées.

En raisonnant comme dans l'exemple 1 on obtient pour la tangente à l'hyperbole l'équation

$$\frac{Xx_0}{a^2}-\frac{Yy_0}{b^2}=1 \quad (|x_0| \geqslant a).$$

Cette tangente coupe l'axe $Ox$ au point d'abscisse $a^2/x_0$ $\left(0 < \frac{a^2}{|x_0|} \leqslant a\right)$, point qui appartient à l'intervalle $]0, a]$ pour $x_0 \geqslant a$ et à l'intervalle $[-a, 0[$ pour $x_0 \leqslant -a$ (fig. 45).

Exemple 3. Trouver l'équation de la tangente à la courbe

$$y^2 = 2px \quad (x \geqslant 0,\ p > 0) \tag{10}$$

en un point $(x_0, y_0)$.

Cette courbe est une *parabole* symétrique par rapport à l'axe $Ox$. Il suffit donc d'étudier la moitié supérieure ($y > 0$). De (10) on déduit que

$$y=\sqrt{2px}. \tag{10'}$$

D'où

$$y'=\frac{p}{\sqrt{2px}}, \quad y_0=\sqrt{2px_0}, \quad y'(x_0)=\frac{p}{\sqrt{2px_0}}=\frac{p}{y_0}.$$

L'équation de la tangente à la parabole en $(x_0, y_0)$ est:

$$Y - y_0 = y'(x_0)(X - x_0)$$

ou

$$Y - y_0 = \frac{p}{y_0}(X - x_0), \quad Yy_0 - y_0^2 = pX - px_0.$$

Comme $y_0^2 = 2px_0$, il vient

$$Yy_0 = p(X + x_0). \tag{11}$$

Donc, pour former l'équation de la tangente à la parabole en un point $(x_0, y_0)$, il faut substituer $Yy_0$ à $y^2$ et $X + x_0$ à $2x$ dans l'équation (10).

La tangente (11) à la parabole (10') en $(x_0, y_0)$ coupe l'axe $Ox$ au point d'abscisse $(-x_0)$ (fig. 46) indépendamment de $p$, autrement

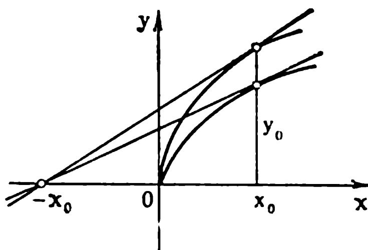



Fig. 46

dit les tangentes aux paraboles $y^2 = 2px$ en $(x_0, \sqrt{2px_0})$ coupent, toutes, l'axe $Ox$ au point $(-x_0)$.

## § 3. Dérivées des fonctions élémentaires

*La constante C.* A toute valeur de $x$ est associée la valeur $y = C$. Donc, à la valeur $x + \Delta x$ correspond la valeur $y + \Delta y = C$. Par suite

$$C' = \lim_{\Delta x \to 0} \frac{C - C}{\Delta x} = \lim_{\Delta x \to 0} \frac{0}{\Delta x} = \lim_{\Delta x \to 0} 0 = 0. \tag{1}$$

*La fonction puissance* $x^n$ $(n = 1, 2, \ldots)$.

$$(x^n)' = nx^{n-1}, \tag{2}$$

car

$$\frac{1}{\Delta x}[(x+\Delta x)^n - x^n] =$$
$$= \frac{1}{\Delta x}\left[x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{2!}x^{n-2}\Delta x^2 + \ldots + \Delta x^n - x^n\right] =$$
$$= nx^{n-1} + \frac{n(n-1)}{2!}x^{n-2}\Delta x + \ldots + \Delta x^{n-1} \underset{\Delta x\to 0}{\longrightarrow} nx^{n-1}.$$

On a les formules

$$(u \pm v)' = u' \pm v', \tag{3}$$
$$(uv)' = uv' + u'v, \tag{4}$$
$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \quad (v \neq 0). \tag{5}$$

On admet que les fonctions $u = u(x)$, $v = v(x)$ sont dérivables en tout point $x$. Dans (5) on admet de plus que $v(x) \neq 0$. On affirme que les premiers membres des égalités (3), (4) et (5) sont dérivables en tout point $x$ et que ces égalités sont réalisées.

En effet, à la valeur $x + \Delta x$ de l'argument sont associées les valeurs $u + \Delta u$ et $v + \Delta v$ des fonctions $u$ et $v$ et

$$\Delta(u \pm v) = [(u+\Delta u) \pm (v + \Delta v)] - (u \pm v) = \Delta u \pm \Delta v,$$
$$(u \pm v)' = \lim_{\Delta x\to 0}\frac{\Delta(u\pm v)}{\Delta x} = \lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x} \pm \lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x} = u' \pm v'.$$

D'autre part,

$$\Delta(uv) = (u+\Delta u)(v+\Delta v) - uv = u\,\Delta v + v\,\Delta u + \Delta u\,\Delta v,$$
$$(uv)' = \lim_{\Delta x\to 0}\frac{\Delta(uv)}{\Delta x} = \lim_{\Delta x\to 0}\frac{u\,\Delta v + v\,\Delta u + \Delta u\,\Delta v}{\Delta x} =$$
$$= u\lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x} + v\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x} + \lim_{\Delta x\to 0}\Delta u \lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x} =$$
$$= uv' + vu' + 0v' = uv' + vu'.$$

Il faut tenir compte du fait que la fonction $u$ est continue, car dérivable, et par suite $\Delta u \to 0$ pour $\Delta x \to 0$.

Enfin

$$\left(\frac{u}{v}\right)' = \lim_{\Delta x\to 0}\left(\frac{u+\Delta u}{v+\Delta v} - \frac{u}{v}\right)\frac{1}{\Delta x} = \lim_{\Delta x\to 0}\frac{v\,\Delta u - u\,\Delta v}{(v+\Delta v)\,v\,\Delta x} =$$
$$= \lim_{\Delta x\to 0}\frac{v\dfrac{\Delta u}{\Delta x} - u\dfrac{\Delta v}{\Delta x}}{(v+\Delta v)\,v} = \frac{u'v - uv'}{v^2}.$$

Ici aussi $\Delta v \to 0$ avec $\Delta x$, car la fonction $v$ est continue, puisqu'elle est dérivable.

Considérons la fonction sin $x$. On a

$$(\sin x)' = \cos x, \tag{6}$$

car

$$(\sin x)' = \lim_{\Delta x \to 0} \frac{\sin(x+\Delta x)-\sin x}{\Delta x} = \lim_{\Delta x \to 0} \frac{2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right)}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}} \lim_{\Delta x \to 0} \cos\left(x+\frac{\Delta x}{2}\right) = 1\cdot\cos x = \cos x$$

(la fonction $\cos x$ étant continue).

On démontre de façon analogue que

$$(\cos x)' = -\sin x, \tag{7}$$

$$(\operatorname{tg} x)' = \sec^2 x = \frac{1}{\cos^2 x}, \tag{8}$$

$$(\operatorname{cotg} x)' = -\operatorname{cosec}^2 x = \frac{-1}{\sin^2 x}. \tag{9}$$

Par exemple,

$$(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos x(\sin x)' - \sin x(\cos x)'}{\cos^2 x} =$$

$$= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

Pour la fonction $y = \log_a x$ $(x > 0)$, on a

$$\frac{\Delta y}{\Delta x} = \frac{\log_a(x+\Delta x) - \log_a x}{\Delta x} = \frac{\log_a\left(1+\frac{\Delta x}{x}\right)}{\Delta x} = \frac{1}{x}\frac{\log_a\left(1+\frac{\Delta x}{x}\right)}{\frac{\Delta x}{x}}.$$

Comme

$$\lim_{u \to 0} \frac{\log_a(1+u)}{u} = \log_a e,$$

il vient

$$(\log_a x)' = \frac{1}{x}\log_a e = \frac{1}{x\ln a}. \tag{10}$$

En particulier,

$$(\ln x)' = \frac{1}{x}. \tag{10'}$$

## § 4. Dérivée d'une fonction composée

THÉORÈME. *Si une fonction $x = \varphi(t)$ est dérivable en un point $t$ et une fonction $y = f(x)$ en un point $x$, alors il en est de même de la fonction composée (par rapport à $t$)*

$$y = F(t) = f[\varphi(t)] \tag{1}$$

*et de plus*

$$F'(t) = f'(x)\,\varphi'(t) \tag{2}$$

*ou*

$$y'_t = y'_x x'_t. \tag{3}$$

DÉMONSTRATION. A un accroissement $\Delta t \neq 0$ de $t$ correspond l'accroissement $\Delta x = \varphi(t + \Delta t) - \varphi(t)$ de $x$. Si $\Delta x \neq 0$, alors

$$\frac{\Delta y}{\Delta t} = \frac{\Delta y}{\Delta x}\frac{\Delta x}{\Delta t}. \tag{4}$$

$\Delta x$ tend vers 0 avec $\Delta t$, car la fonction $x = \varphi(t)$ est continue en $t$ (puisqu'elle est dérivable). Donc

$$y'_t = \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = y'_x x'_t \tag{5}$$

et la formule (3) est démontrée sous réserve que $\Delta x \neq 0$. Plus exactement sous réserve qu'existe un $\delta > 0$ tel que $|\Delta x| > 0$ pour tous les $\Delta t$ tels que $0 < |\Delta t| < \delta$.

Si cette condition n'est pas remplie, on démontre la formule (3) de la manière suivante. Si la condition indiquée n'est pas réalisée, il existe une suite $\{\Delta t_k\}$ convergeant vers 0, telle que $\Delta x_k = 0$ pour tous les $k = 1, 2, \ldots$, et alors

$$\lim_{\Delta t_k \to 0} \frac{\Delta x_k}{\Delta t_k} = \lim_{\Delta t_k \to 0} \frac{0}{\Delta t_k} = 0. \tag{6}$$

Ceci montre que $x'_t = 0$, car la limite de $\frac{\Delta x_k}{\Delta t_k}$ ne dépend pas du mode de tendance de $\Delta t_k$ vers 0. (En effet, on a admis l'existence de $x'_t$.) Considérons deux sortes de suites $\{\Delta t_k\}$ convergeant vers 0. La première sorte est composée de suites telles que $\Delta x_k \neq 0$ pour tous les $k = 1, 2, \ldots$ Pour ces suites on a par analogie avec (5):

$$\lim_{\Delta t_k \to 0} \frac{\Delta y_k}{\Delta t_k} = \lim_{\Delta x_k \to 0} \frac{\Delta y_k}{\Delta x_k} \lim_{\Delta t_k \to 0} \frac{\Delta x_k}{\Delta t_k} = y'_x x'_t = 0. \tag{7}$$

Les suites $\{\Delta t_k\}$ de la deuxième sorte sont telles que $\Delta x_k = 0$ $(k = 1, 2, \ldots)$. Alors

$$\Delta y_k = F(t + \Delta t_k) - F(t) = f(x + \Delta x_k) - f(x) = f(x) - f(x) = 0$$
$$(k = 1, 2, \ldots)$$

et

$$\lim_{\Delta t_k \to 0} \frac{\Delta y_k}{\Delta t_k} = \lim_{\Delta t_k \to 0} \frac{0}{\Delta t_k} = 0. \tag{8}$$

Si maintenant est donnée une suite $\{\Delta t_k\}$ convergeant vers 0, on en extrait les $\Delta t_k$ pour lesquels $\Delta x_k \neq 0$. Si ces $\Delta t_k$ sont en nombre infini, ils forment une suite de la première sorte pour laquelle (7) a été démontrée. Les $\Delta t_k$ restants peuvent former une suite, mais ce sera une suite de la deuxième sorte pour laquelle (8) a été démontrée. Donc, pour toute suite $\{\Delta t_k\}$ telle que $\Delta t_k \to 0$, on a

$$\lim_{\Delta t_k \to 0} \frac{\Delta y_k}{\Delta t_k} = 0 = y'_x \cdot x'_t,$$

ce qui prouve entièrement le théorème.

La formule (1) peut prendre une forme plus compliquée. Si par exemple $z = f(y)$, $y = \varphi(x)$, $x = \psi(\xi)$ sont dérivables aux points correspondants, alors $z'_\xi = z'_y y'_x x'_\xi$.

Exemple 1. $y = \ln \sin^2 x \ (x \neq k\pi)$.

Posons $y = \ln u$, $u = v^2$, $v = \sin x$. Il vient

$$y'_x = y'_u u'_v v'_x = \frac{1}{u} 2v \cos x = \frac{2 \sin x \cos x}{\sin^2 x} = 2 \operatorname{cotg} x.$$

Exemple 2. $y = \sin(x^2 + 2x - 1)$.

Posons $u = x^2 + 2x - 1$. Alors

$$y'_x = y'_u \cdot u'_x = \cos u \cdot (2x + 2) = 2(x + 1) \cos(x^2 + 2x - 1).$$

Dans les calculs, les variables auxiliaires $u$, $v$, ... ne sont pas généralement introduites mais seulement sous-entendues.

Dans le cas de l'exemple 1, les calculs sont conduits de la manière suivante:

$$y'_x = \frac{1}{\sin^2 x} (\sin^2 x)' = \frac{1}{\sin^2 x} 2 \sin x (\sin x)' = \frac{2}{\sin x} \cos x = 2 \operatorname{cotg} x.$$

Ou de façon plus concise

$$y'_x = \frac{1}{\sin^2 x} 2 \sin x \cos x = 2 \operatorname{cotg} x.$$

## § 5. Dérivée de la réciproque d'une fonction

Théorème. *Soit $y = f(x)$ une fonction strictement croissante, continue sur un intervalle $]a, b[$ et admettant une dérivée $f'(x)$ non nulle en un point $x \in ]a, b[$. Alors la fonction $x = f^{-1}(y) = g(y)$ réciproque de $f$ admet aussi une dérivée au point correspondant, définie par*

$$g'(y) = \frac{1}{f'(x)} \tag{1}$$

*ou*

$$x'_y = \frac{1}{y'_x}. \tag{1'}$$

DÉMONSTRATION. On sait que la fonction réciproque $x = g(y)$ est strictement croissante et continue sur l'intervalle $]A, B[$, où

$$A = \inf_{x \in ]a, b[} f(x), \quad B = \sup_{x \in ]a, b[} f(x)$$

(voir théorème 1′, § 6, chap. 3).

Donnons à $y$ un accroissement $\Delta y \neq 0$. A $\Delta y$ correspond un accroissement $\Delta x$ de la fonction réciproque, un accroissement non nul puisque $f$ est strictement croissante. Donc

$$\frac{\Delta x}{\Delta y} = \frac{1}{\frac{\Delta y}{\Delta x}}.$$

Si maintenant $\Delta y \to 0$, alors $\Delta x \to 0$, puisque $g(y)$ est continue; mais $\frac{\Delta y}{\Delta x} \to f'(x) \neq 0$ pour $\Delta x \to 0$, donc existe la limite

$$\lim_{\Delta y \to 0} \frac{\Delta x}{\Delta y} = \frac{1}{\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}} = \frac{1}{f'(x)}.$$

Ceci prouve (1).

REMARQUE. *Si $f'(x) \neq 0$ est continue sur $]a, b[$, il en est de même de $g'(y)$ sur $]A, B[$.*

Ceci résulte de (1), où l'on peut poser $x = g(y)$:

$$g'(y) = \frac{1}{f'[g(y)]}, \quad y \in ]A, B[.$$

En effet, la fonction $f'[g(y)]$ est continue, car elle est la composée de fonctions continues $f'$ et $g$.

## § 6. Dérivées des fonctions élémentaires (suite)

1. $y = a^x$. Sa réciproque est la fonction $x = \log_a y$. Donc

$$y'_x = \frac{1}{x'_y} = \frac{1}{\frac{1}{y \ln a}} = y \ln a = a^x \ln a, \quad \text{i.e.} \ (a^x)' = a^x \ln a.$$

En particulier

$$(e^x)' = e^x, \quad (e^{-x})' = -e^{-x}.$$

2. $y = \operatorname{Arcsin} x$ ($|x| < 1$, $-\pi/2 < y < \pi/2$). $x = \sin y$ étant la fonction réciproque, on a

$$y'_x = \frac{1}{x'_y} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}},$$

c'est-à-dire que

$$(\operatorname{Arc} \sin x)' = \frac{1}{\sqrt{1-x^2}}.$$

La racine carrée est prise avec le signe $+$, car $\cos y > 0$ sur $]-\pi/2, \pi/2[$.

3.

$$(\operatorname{Arc}\cos x)' = \left(\frac{\pi}{2} - \operatorname{Arc}\sin x\right)' = -\frac{1}{\sqrt{1-x^2}}.$$

4. $y = \operatorname{Arc}\operatorname{tg} x$. Sa réciproque est $x = \operatorname{tg} y$ $(-\infty < x < \infty, -\pi/2 < y < \pi/2)$. Donc

$$(\operatorname{Arc}\operatorname{tg} x)' = \frac{1}{(\operatorname{tg} y)'} = \cos^2 y = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2},$$

i.e.

$$(\operatorname{Arc}\operatorname{tg} x)' = \frac{1}{1+x^2}.$$

5. On démontre de façon analogue que

$$(\operatorname{Arc}\operatorname{cotg} x)' = -\frac{1}{1+x^2}.$$

6. *Dérivée de la fonction puissance* $x^\alpha$ ($x > 0$, $\alpha$ réel). On a

$$x^\alpha = e^{\alpha \ln x}.$$

Les fonctions $e^u$ et $\alpha \ln x$ étant dérivables, le théorème de la dérivée d'une fonction composée nous donne

$$(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x} \cdot \frac{\alpha}{x} = \alpha \frac{x^\alpha}{x} = \alpha x^{\alpha-1}.$$

Donc

$$(x^\alpha)' = \alpha x^{\alpha-1}.$$

Ce résultat est en accord avec la formule (2) du § 3 de la dérivée de la fonction $x^n$ ($x \in ]-\infty, \infty[$), où $n$ est un entier naturel.

7. *La fonction* $y = u(x)^{v(x)}$ ($u > 0$). Si $u(x)$ et $v(x)$ sont dérivables, il en est de même de la fonction

$$u^v = e^{v \ln u} \tag{1}$$

et

$$(u^v)' = e^{v \ln u}(v \ln u)' = u^v\left(\frac{v}{u} u' + v' \ln u\right). \tag{2}$$

L'expression

$$[\ln f(x)]' = \frac{f'(x)}{f(x)} \tag{3}$$

s'appelle *dérivée logarithmique de la fonction* $f$.

Comme

$$\ln u^v = v \ln u,$$

on a en vertu de (3)

$$\frac{(u^v)'}{u^v} = (v \ln u)' = v' \ln u + \frac{vu'}{u};$$

d'où (2).

8. *Les fonctions hyperboliques*.

$$(\mathrm{sh}\, x)' = \left(\frac{e^x - e^{-x}}{2}\right)' = \frac{e^x + e^{-x}}{2} = \mathrm{ch}\, x,$$

$$(\mathrm{ch}\, x)' = \left(\frac{e^x + e^{-x}}{2}\right)' = \frac{e^x - e^{-x}}{2} = \mathrm{sh}\, x,$$

$$(\mathrm{th}\, x)' = \left(\frac{\mathrm{sh}\, x}{\mathrm{ch}\, x}\right)' = \frac{\mathrm{ch}^2 x - \mathrm{sh}^2 x}{\mathrm{ch}^2 x} = \frac{1}{\mathrm{ch}^2 x},$$

$$(\mathrm{coth}\, x)' = \left(\frac{\mathrm{ch}\, x}{\mathrm{sh}\, x}\right)' = \frac{\mathrm{sh}^2 x - \mathrm{ch}^2 x}{\mathrm{sh}^2 x} = \frac{-1}{\mathrm{sh}^2 x} \quad (x \neq 0).$$

9. $y = \mathrm{Arg\,sh}\, x$. Sa réciproque étant $x = \mathrm{sh}\, y$, on a

$$(\mathrm{Arg\,sh}\, x)' = \frac{1}{(\mathrm{sh}\, y)'} = \frac{1}{\mathrm{ch}\, y} = \frac{1}{\sqrt{1+\mathrm{sh}^2 y}} = \frac{1}{\sqrt{1+x^2}}$$

(voir plus bas l'exemple 2 du § 12).

## § 7. Différentielle d'une fonction

Supposons qu'une fonction $f$ admet une dérivée (finie) en un point $x$:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x).$$

Alors pour $\Delta x$ assez petit, on peut mettre $\frac{\Delta y}{\Delta x}$ sous la forme d'une somme de $f'(x)$ et d'une fonction, que l'on désignera par $\varepsilon(\Delta x)$, qui tend vers 0 avec $\Delta x$:

$$\frac{\Delta y}{\Delta x} = f'(x) + \varepsilon(\Delta x) \quad (\varepsilon(\Delta x) \to 0, \quad \Delta x \to 0).$$

L'accroissement de $f$ en $x$ s'écrit

$$\Delta y = f'(x)\, \Delta x + \Delta x \cdot \varepsilon(\Delta x) \quad (\varepsilon(\Delta x) \to 0, \quad \Delta x \to 0)$$

ou

$$\Delta y = f'(x)\, \Delta x + \underset{\Delta x \to 0}{o(\Delta x)}. \tag{1}$$

Rappelons que l'expression $\underset{\Delta x \to 0}{o(\Delta x)}$ doit être comprise comme une fonction de $\Delta x$, dont le quotient par $\Delta x$ tend vers 0 avec $\Delta x$.

DÉFINITION. *On dit qu'une fonction $f$ est différentiable en un point $x$ si son accroissement $\Delta y$ en $x$ peut être mis sous la forme*

$$\Delta y = A \cdot \Delta x + \underset{\Delta x \to 0}{o(\Delta x)}, \tag{2}$$

*où $A$ dépend de $x$ mais pas de $\Delta x$.*

THÉORÈME. *Pour qu'une fonction $f$ soit différentiable en un point $x$ il est nécessaire et suffisant qu'elle admette une dérivée finie en ce point. Ceci étant, $A = f'(x)$.*

Donc, dire que $f$ admet une dérivée en un point $x$ revient à dire que $f$ est différentiable en $x$.

DÉMONSTRATION. La condition suffisante a été démontrée plus haut : l'existence de la dérivée finie $f'(x)$ nous a permis de représenter $\Delta y$ sous la forme (1). Reste à poser $f'(x) = A$.

*Nécessité.* Supposons que la fonction $f$ est différentiable en $x$. Alors de (2) et en admettant que $\Delta x \neq 0$, on obtient

$$\frac{\Delta y}{\Delta x} = A + \frac{\underset{\Delta x \to 0}{o(\Delta x)}}{\Delta x} = A + \underset{\Delta x \to 0}{o(1)}.$$

Le second membre converge vers $A$ pour $\Delta x \to 0$. Donc

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = A$$

ou encore

$$f'(x) = A.$$

Soit $y = f(x)$ une fonction différentiable en un point $x$. Son accroissement $\Delta y$ est la somme de deux termes. Le premier, $A\,\Delta x$, est proportionnel à $\Delta x$; on dit encore que c'est une fonction homogène linéaire de $\Delta x$. Le second, $\underset{\Delta x \to 0}{o(\Delta x)}$, est un infiniment petit d'ordre supérieur à $\Delta x$. Si $A \neq 0$, le second terme tend vers 0 plus vite que le premier pour $\Delta x \to 0$. C'est pourquoi le premier terme $A\,\Delta x = f'(x)\,\Delta x$ s'appelle *terme principal de l'accroissement* $\Delta y$ (pour $\Delta x \to 0$. Voir définition de la fin du § 10, chap. 3). On l'appelle aussi *différentielle de la fonction $y$* et on le note $dy$:

$$dy = df = f'(x)\,\Delta x.$$

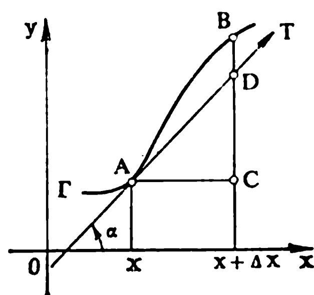


Fig. 47

La figure 47 représente le graphe $\Gamma$ d'une fonction $y = f(x)$ : $T$ est la tangente à $\Gamma$ en un point $A$ d'abscisse $x$; $f'(x) = \operatorname{tg}\alpha$, où $\alpha$ est l'angle de la tangente $T$ avec l'axe $Ox$:

$$dy = f'(x)\Delta x = \operatorname{tg}\alpha\Delta x = CD, \quad DB = \Delta y - dy = \underset{\Delta x \to 0}{o(\Delta x)}.$$

Donc, *la différentielle de la fonction $y$ en $x$ est l'accroissement de l'ordonnée d'un point situé sur la tangente* ($dy = CD$).

En général, $dy \neq \Delta y$, car $\Delta y = dy + \underset{\Delta x \to 0}{o}(\Delta x)$ et le second membre n'est pas toujours nul. Seule la fonction linéaire $y = Ax + B$ est telle que $\Delta y = A\Delta x = dy$ pour tout $x$. En particulier, pour $y = x$, $dy = dx = \Delta x$, c'est-à-dire que *la différentielle et l'accroissement de la variable indépendante sont égaux* ($dx = \Delta x$). C'est la raison pour laquelle la différentielle d'une fonction $f$ se note habituellement :

$$dy = f'(x)\,dx,$$

d'où

$$f'(x) = \frac{dy}{dx},$$

c'est-à-dire que *la dérivée de la fonction $f$ en un point $x$ est égale au quotient de la différentielle de la fonction en ce point par la différentielle de la variable indépendante $x$.*

Ceci explique que l'on désigne la dérivée par $dy/dx$.

Signalons que *la différentielle $dx$ de la variable indépendante ne dépend pas de $x$, elle est égale à l'accroissement $\Delta x$ de $x$. Au contraire, la différentielle $dy$ d'une fonction $y$ (différente de $x$) dépend de $x$ et de $dx$.*

On a les formules

$$d(u \pm v) = du \pm dv, \tag{3}$$

$$d(u \cdot v) = u\,dv + v\,du, \tag{4}$$

$$d(cu) = c\,du \quad (c \text{ est une constante}), \tag{5}$$

$$d\left(\frac{u}{v}\right) = \frac{v\,du - u\,dv}{v^2} \quad (v \neq 0), \tag{6}$$

où l'on admet que $u$ et $v$ sont des fonctions différentiables en $x$.

La formule (6) se démontre comme suit :

$$d\left(\frac{u}{v}\right) = \left(\frac{u}{v}\right)' dx = \frac{vu'\,dx - uv'\,dx}{v^2} = \frac{v\,du - u\,dv}{v^2}.$$

Si une fonction $y = f(x)$ est différentiable en un point $x$, en vertu de la formule (1) on peut mettre son accroissement $\Delta y$ sous la forme :

$$\Delta y = dy + \underset{\Delta x \to 0}{o}(\Delta x).$$

De là il s'ensuit que la différentielle d'une fonction peut servir, pour $\Delta x$ assez petit, de bonne approximation de l'accroissement de cette fonction. On a alors l'égalité approchée

$$\Delta y \approx dy = f'(x)\,dx, \tag{7}$$

qui est d'un usage très courant.

EXEMPLE. Si l'on admet que

$$\sqrt[3]{8{,}001} \approx \sqrt[3]{8} = 2,$$

alors l'erreur est approximativement égale à la différentielle de la fonction $y = x^{1/3}$ au point $x = 8$ pour $\Delta x = 0{,}001$:

$$dy = \frac{1}{3} x^{-2/3} \Delta x = \frac{1}{3} 8^{-2/3} \cdot 0{,}001 = 1/12\,000.$$

Les méthodes qui seront exposées au § 14 nous permettront de juger de l'exactitude de nos calculs.

## § 8. Autre définition de la tangente

Si la dérivée $f'(x_0)$ est finie, on peut la définir d'une autre manière équivalente.

Soient $\Gamma$ la courbe représentative d'une fonction $y = f(x)$, $A = (x_0, f(x_0))$ un point de $\Gamma$. Considérons une droite $L$ passant par $A$. Elle a pour équation $y - y_0 = m(x - x_0)$. Soit $B = (x, f(x))$ un point de $\Gamma$ non confondu avec $A$. La distance de $B$ à $L$ dans la direction de l'axe $Oy$ est égale à

$$\rho(x) = |f(x) - y_0 - m(x - x_0)|. \quad (1)$$

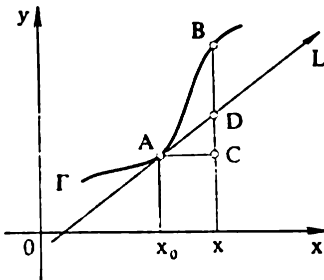


Fig. 48

Sur la figure 48, $\rho(x) = BD$.

On dit que *la droite $L$ est tangente à $\Gamma$ en $A$ si*

$$\rho(x) \underset{x \to x_0}{=} o(x - x_0). \quad (2)$$

Si la droite $L$ est tangente à $\Gamma$ en $A$ au sens de la première définition, alors $m = f'(x_0)$. Comme $f$ est dérivable, il vient

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + o(x - x_0), \quad x \to x_0,$$

d'où

$$\rho(x) = |f(x) - f(x_0) - f'(x_0)(x - x_0)| = o(x - x_0), \quad x \to x_0,$$

autrement dit, la droite $L$ est tangente à la courbe au sens de la deuxième définition.

Réciproquement, si $L$ est tangente au sens de la deuxième définition, alors (voir (1) et (2))

$$\rho(x) = |f(x) - f(x_0) - m(x - x_0)| = o(x - x_0), \quad x \to x_0,$$

ou ce qui revient au même

$$f(x) - f(x_0) = m(x - x_0) + o(x - x_0) \quad \text{pour} \quad x \to x_0.$$

Ceci montre que la fonction $f$ est dérivable en $x_0$ et $m = f'(x_0)$. Donc, $L$ est tangente au sens de la première définition et a pour équation

$$y - y_0 = f'(x_0)(x - x_0).$$

REMARQUE. De ce qui précède il s'ensuit que la courbe d'équation $y = f(x)$ admet une tangente en un point $(x_0, f(x_0))$ si et seulement si la fonction $f$ est dérivable en $x_0$.

## § 9. Dérivée d'ordre supérieur

Soit donnée une fonction $f$ sur un intervalle $]a, b[$. Si elle est dérivable sur $]a, b[$, sa dérivée $f'(x)$ sera dite *dérivée première*. Si la dérivée première est à son tour dérivable sur $]a, b[$, sa dérivée est dite *dérivée seconde* ou *d'ordre deux de* $f$ et se note:

$$f''(x) = f^{(2)}(x) = (f'(x))' \quad \text{ou} \quad y'' = (y')'.$$

D'une façon générale, *on appelle dérivée d'ordre n d'une fonction f la dérivée première de la dérivée d'ordre n — 1 de f* et l'on note:

$$f^{(n)}(x) = (f^{(n-1)}(x))' \text{ ou encore } y^{(n)} = (y^{(n-1)})'.$$

Si le point $x$ est fixé, le symbole $f^{(n)}(x)$ désigne la dérivée d'ordre $n$ de $f$ en $x$. Son existence implique celle de la dérivée $f^{(n-1)}$ en $x$ et dans un voisinage de $x$.

EXEMPLES.

1. $(e^x)^{(n)} = e^x$.

2. $(a^x)' = a^x \ln a$, $(a^x)'' = a^x \ln^2 a$, $\ldots$, $(a^x)^{(n)} = a^x \ln^n a$.

3. $(x^m)' = mx^{m-1}$, $(x^m)'' = m(m-1)x^{m-2}$, $\ldots$, $(x^m)^{(n)} = m(m-1)\ldots(m-n+1)x^{m-n}$.

Si $m$ est un entier naturel, il est évident que

$$(x^m)^{(m)} = m! \text{ et } (x^m)^{(n)} = 0 \quad (n > m).$$

4. $(\sin x)' = \cos x = \sin\left(x + \frac{\pi}{2}\right)$,

$$(\sin x)'' = \left[\sin\left(x+\frac{\pi}{2}\right)\right]' = \sin\left(x + 2\frac{\pi}{2}\right),$$

. . . . . . . . . . . . . . . . . . . . . . .

$$(\sin x)^{(n)} = \sin\left(x + n\frac{\pi}{2}\right).$$

5. $(\cos x)^{(n)} = \cos\left(x + n\frac{\pi}{2}\right)$.

A noter toutefois qu'il n'est pas toujours possible d'exprimer les dérivées d'une fonction quelconque par une formule générale.

EXERCICE. Etablir par récurrence la formule (de Leibniz) de la dérivée d'ordre $n$ du produit de deux fonctions:

$$(uv)^{(n)} = \sum_{k=0}^{n} C_n^k u^{(n-k)} v^{(k)},$$

où $u$ et $v$ sont des fonctions dérivables jusqu'à l'ordre $n$ compris et

$$C_n^k = \frac{n(n-1)\ldots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}, \quad 0! = 1.$$

## § 10. Différentielle d'ordre supérieur. Invariance de la différentielle première

Etant donné une fonction $y = f(x)$ sur un intervalle $]a, b[$, on peut de toute évidence la représenter d'une infinité de manières comme une fonction composée:

$$y = \varphi(z), \quad z = \psi(x).$$

Donc, $y$ peut être traitée comme une fonction de $x$ ($y = f(x)$) et comme une fonction de $z$ ($y = \varphi(z)$), où $z$ est à son tour une fonction de $x$ ($z = \psi(x)$).

La variable $x$ sera dite *indépendante*, car elle ne dépendra d'aucune autre. La variable $z$ sera dite *dépendante* (de $x$).

La différentielle de la fonction $y = f(x)$ en un point $x$ est, nous le savons, le produit de la dérivée de $f$ en ce point par la différentielle de la variable indépendante:

$$dy = f'(x)\, dx.$$

Ici $dx$ est un nombre arbitraire qui *ne dépend pas de* $x$. Donc

$$(dx)' = 0.$$

La différentielle d'une fonction s'appelle encore *différentielle première*.

Par définition, *la différentielle seconde d'une fonction* $y = f(x)$ *en un point* $x$ *est la différentielle de la différentielle première en ce point* et se note:

$$d^2y = d(dy).$$

Pour calculer la différentielle seconde, il faut dériver le produit $f'(x)\, dx = dy$ par rapport à $x$ en traitant $dx$ comme une constante (ne dépendant pas de $x$) et multiplier le résultat par $dx$:

$$d^2y = d[f'(x)\, dx] = dx\, d[f'(x)] = f''(x)\, dx^2.$$

D'une façon générale, *on appelle différentielle d'ordre* $n$ *d'une fonction* $y = f(x)$ *la différentielle première de la différentielle d'ordre* $(n-1)$ *de cette fonction* et on note:

$$d^ny = d(d^{n-1}y).$$

Il est évident que

$$d^ny = f^{(n)}(x)\, dx^n. \qquad (1)$$

**En effet, cette formule étant vraie pour $n = 1$, si l'on admet qu'elle l'est pour $n - 1$, on obtient**

$$d^n y = d\,[f^{(n-1)}(x)\,dx^{n-1}] = dx^{n-1} d\,[f^{(n-1)}(x)] = f^{(n)}(x)\,dx^n.$$

Il est évident que l'existence de la différentielle d'ordre $n$ de la fonction $y = f(x)$ en un point $x$ implique nécessairement celle de la dérivée $f^{(n)}(x)$ d'ordre $n$ en ce point.

On a en vertu de (1)

$$y_x^{(n)} = f^{(n)}(x) = \frac{d^n y}{dx^n}, \tag{2}$$

autrement dit, *la dérivée d'ordre n de la fonction y par rapport à x est égale au quotient de la différentielle d'ordre n de y par* $dx^n = (dx)^n$.

Nous verrons dans la suite que la formule (2) est inexacte si on y remplace la variable indépendante $x$ par la variable dépendante $z$ (voir formule (4) plus bas).

Nous avons défini les différentielles d'une fonction $y = f(x)$, où $x$ est une *variable indépendante*. Or, nous avons vu que la fonction $y$ peut encore s'écrire sous la forme

$$y = \varphi(z),$$

où $z$ est une fonction de $x$ ($z = \psi(x)$, $f(x) = \varphi[\psi(x)]$). Comment dans ce cas exprimer les différentielles dans le langage de la variable (dépendante) $z$.

Pour la différentielle première on a

$$dy = y'_x dx = y'_z z'_x\, dx = y'_z\,(z'_x dx) = y'_z dz.$$

Nous remarquons que la différentielle de la fonction $y$ est égale au produit de sa dérivée $y'_z$ par $dz$:

$$dy = y'_z\, dz, \tag{3}$$

c'est-à-dire que *la différentielle première de la fonction y s'exprime par la même formule, qu'elle soit traitée comme une fonction de la variable indépendante x ou de la variable dépendante z.*

La forme de la différentielle première étant préservée (voir (3)) on dit qu'elle est *invariante*.

La situation est différente pour les différentielles d'ordre supérieur. En effet, si l'on considère que $y$ est une fonction de $z$ ($y = \varphi(z)$), on obtient (cf. § 7, (6))

$$d^2 y = d\,(dy) = d\,[\varphi'(z)\,dz] = dz\,d\,(\varphi'(z)) + \varphi'(z)\,d\,(dz) = \\ = \varphi''(z)\,dz^2 + \varphi'(z)\,d^2 z. \tag{4}$$

Dans la dernière égalité on s'est servi de l'invariance de la différentielle première en vertu de laquelle $d(\varphi'(z)) = \varphi''(z)\,dz$, ainsi que du fait que $d(dz) = d^2z$. Il ne faut pas négliger la quantité $d^2z = \psi''(x)\,dx^2$. En effet, elle n'est nulle (pour tous les $x$) que si $\psi(x)$ est une fonction linéaire ($\psi(x) = Ax + B$).

Nous constatons qu'exprimée en fonction de $z$, la différentielle seconde change de forme : au nombre $\varphi''(z)\,dz^2$ s'est ajouté le terme $\varphi'(z)\,d^2z$ qui n'est pas nul en général.

## § 11. Dérivation de fonctions données sous forme paramétrique

Supposons que la dépendance de $y$ par rapport à $x$ est exprimée au moyen d'un paramètre $t$:

$$\begin{cases} x = \varphi(t), \\ y = \psi(t), \end{cases} \quad t \in ]a, b[. \tag{1}$$

Cela signifie que la fonction $x = \varphi(t)$ admet une fonction réciproque et qu'on peut représenter explicitement $y$ en fonction de $x$:

$$y = \psi[\varphi^{-1}(x)]. \tag{2}$$

On cherchera la dérivée de $y$ par rapport à $x$ en fonction des dérivées de $x$ et de $y$ par rapport à $t$. La différentielle première étant invariante, on a $y'_x = dy/dx$. Or, $dy = y'_t dt$, $dx = x'_t dt$. Donc

$$y'_x = \frac{y'_t}{x'_t} \quad (x'_t \neq 0). \tag{3}$$

Pour la dérivée seconde, on obtient

$$y''_x = \frac{d}{dx}\, y'_x = \frac{d}{dx}\left(\frac{y'_t}{x'_t}\right) = \frac{d}{dt}\left(\frac{y'_t}{x'_t}\right)\frac{dt}{dx} = \frac{x'_t y''_t - y'_t x''_t}{(x'_t)^3}. \tag{4}$$

On obtiendrait par analogie les dérivées $y_x^{(n)}$ d'ordre $n > 2$ en fonction des dérivées de $x$ et de $y$ par rapport à $t$.

## § 12. Théorèmes de la moyenne

Par définition, *une fonction $f$ a un maximum (resp. minimum) local en un point $x = c$ s'il existe un voisinage $V(c) = ]c - \delta, c + \delta[$ dans lequel*

$$f(c) \geqslant f(x), \qquad \forall x \in V(c) \tag{1}$$

$$(resp.\ f(x) \geqslant f(c), \qquad \forall x \in V(c)). \tag{1'}$$

On appelle *extrémum local* un maximum ou un minimum local.

REMARQUE 1. Si une fonction $f$ est continue sur un intervalle fermé $[a, b]$ et a son maximum (resp. minimum) en un point $c \in ]a, b[$,

il est évident qu'elle a en même temps un maximum (resp. minimum) local en $c$. Si par contre $f$ a son maximum (resp. minimum) en une extrémité de l'intervalle $[a, b]$, alors elle n'a pas un maximum (resp. minimum) local en cette extrémité, car elle n'est pas définie à droite ou à gauche de cette extrémité.

La figure 49 représente le graphe d'une fonction $y = f(x)$ continue sur $[a, b]$. La fonction $f$ a un minimum local aux points $x_2$ et $x_4$ et un maximum local aux points $x_1$ et $x_3$. A noter qu'en $a$ et en $b$ la fonction $f$ a respectivement un minimum et un maximum locaux unilatéraux.

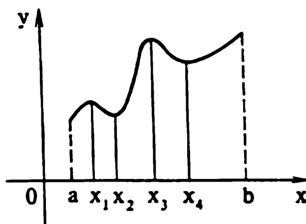


Fig. 49

THÉORÈME 1 (FERMAT[1]). *Si une fonction $f$ est dérivable en un point $c$ et a en ce point un extrémum local, alors $f'(c) = 0$.*

DÉMONSTRATION. Pour fixer les idées on admettra que $f$ a un maximum local en $c$. Par définition de la dérivée on a

$$f'(c) = \lim_{\Delta x \to 0} \frac{f(c+\Delta x) - f(c)}{\Delta x}.$$

Comme $f(c) \geqslant f(x)$, $\forall x \in V(c)$, on a pour les $\Delta x > 0$ assez petits

$$\frac{f(c+\Delta x) - f(c)}{\Delta x} \leqslant 0,$$

d'où en passant à la limite pour $\Delta x \to 0$

$$f'(c) \leqslant 0. \tag{2}$$

Si $\Delta x < 0$, alors

$$\frac{f(c+\Delta x) - f(c)}{\Delta x} \geqslant 0,$$

donc, en passant à la limite pour $\Delta x \to 0$, on obtient

$$f'(c) \geqslant 0. \tag{3}$$

Des relations (2) et (3) il s'ensuit que $f'(c) = 0$.

THÉORÈME 2 (ROLLE[2]). *Si une fonction $y = f(x)$ est continue sur $[a, b]$, dérivable sur $]a, b[$ et $f(a) = f(b)$, il existe alors un point $\xi \in$ $\in ]a, b[$ tel que $f'(\xi) = 0$.*

DÉMONSTRATION. Si $f$ est constante sur $[a, b]$, alors $f'(\xi) = 0$ pour tous les $\xi \in ]a, b[$.

Supposons maintenant que $f$ n'est pas constante sur $[a, b]$. Comme $f$ est continue sur $[a, b]$, il existe un point $x_1 \in [a, b]$ en lequel $f$ a un maximum (cf. chap. 3, § 5, théorème 2) et un point $x_2 \in [a, b]$ en

[1]) Pierre de Fermat, mathématicien français (1601-1665).
[2]) Michel Rolle, mathématicien français (1652-1719), à qui l'on doit la démonstration de ce théorème pour les polynômes.

lequel elle a un minimum. Ces points ne sont pas confondus avec les extrémités $a$ et $b$ de $[a, b]$, sinon on aurait

$$\max_{x\in[a,b]} f(x) = \min_{x\in[a,b]} f(x) = f(a) = f(b)$$

et $f$ serait constante sur $[a, b]$. Donc, l'un des points $x_1$ et $x_2$ est compris dans l'intervalle $]a, b[$. Désignons-le par $\xi$. La fonction $f$ a un extrémum local en $\xi$, de plus elle est dérivable en ce point, puisqu'on a admis qu'elle l'était sur $]a, b[$. Donc, d'après le théorème de Fermat, on a $f'(\xi) = 0$.

REMARQUE 2. Le théorème de Rolle est valable même pour un intervalle ouvert $]a, b[$, pourvu que

$$\lim_{\substack{x\to a\\ x>a}} f(x) = \lim_{\substack{x\to b\\ x<b}} f(x).$$

REMARQUE 3. Le théorème de Rolle est mis en défaut si $f'(x)$ n'existe pas en un point au moins de $]a, b[$. Exemple: $y = |x|$ sur $[-1, 1]$. Dans l'énoncé du théorème on ne peut pas non plus remplacer la continuité sur $[a, b]$ par la continuité sur $]a, b[$. Exemple:

$$y = \begin{cases} 1, & x = 0, \\ x, & 0 < x \leqslant 1, \end{cases}$$

pour laquelle le point $x = 0$ est un point de discontinuité.

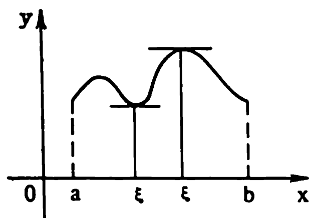


Fig. 50

REMARQUE 4. Le théorème de Rolle admet une interprétation géométrique simple. Si les conditions du théorème sont remplies, alors le graphe de la fonction $y = f(x)$ (fig. 50) présente un point en lequel la tangente est parallèle à l'axe $Ox$.

THÉORÈME 3 (CAUCHY). *Si des fonctions $f(x)$ et $g(x)$ sont continues sur $[a, b]$, dérivables sur $]a, b[$ et $g'(x) \neq 0$ sur $]a, b[$, il existe alors un point $\xi \in ]a, b[$ tel que*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}. \tag{4}$$

DÉMONSTRATION. On remarquera que $g(b) - g(a) \neq 0$, sinon il existerait d'après le théorème de Rolle un point $\xi$ tel que $g'(\xi) = 0$, ce qui est contraire à l'hypothèse. Considérons la fonction auxiliaire

$$F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}[g(x) - g(a)].$$

La fonction $F$ est par hypothèse continue sur $[a, b]$, dérivable sur $]a, b[$ et $F(a) = 0$, $F(b) = 0$. Le théorème de Rolle nous dit

qu'il existe alors un point $\xi \in ]a, b[$ tel que $F'(\xi) = 0$. Or,

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x),$$

donc, en substituant $\xi$ à $x$ on retrouve la relation (4). C.q.f.d.

REMARQUE 5. Dans la formule (4) de Cauchy on voit qu'il n'est pas nécessaire que $a < b$.

Une conséquence du théorème de Cauchy pour $g(x) = x$ est le théorème de Lagrange.

THÉORÈME 4 (DE LA MOYENNE DE LAGRANGE [1]). *Supposons qu'une fonction $f(x)$ est continue sur un intervalle $[a, b]$ et a une dérivée sur $]a, b[$. Il existe alors un point $c \in ]a, b[$ en lequel*

$$f(b) - f(a) = (b - a)f'(c). \qquad (5)$$

Le théorème de Lagrange admet une interprétation géométrique simple. En effet, si l'on écrit la relation (5) sous la forme

$$\frac{f(b) - f(a)}{b - a} = f'(c) \quad (a < c < b),$$

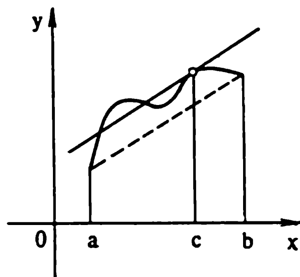


Fig. 51

on constate que le premier membre est la pente de la corde passant par les points $(a, f(a))$ et $(b, f(b))$ du graphe de la fonction $y = f(x)$, et le second membre, la pente de la tangente à cette courbe en un point d'abscisse $c \in ]a, b[$. Le théorème de Lagrange dit que si une courbe (fig. 51) est le graphe d'une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$, alors cette courbe présente un point d'abscisse $c \in ]a, b[$ en lequel la tangente est parallèle à la corde passant par les points $(a, f(a))$ et $(b, f(b))$.

La relation (5) s'appelle *formule (de Lagrange) des accroissements finis*. Si l'on représente la valeur intermédiaire $c$ par

$$c = a + \theta(b - a),$$

où $\theta$ est un nombre tel que $\theta \in ]0, 1[$, alors la formule de Lagrange devient

$$f(b) - f(a) = (b - a) f'(a + \theta(b - a)). \qquad (6)$$

Cette formule est valable visiblement pour $a \geqslant b$.

THÉORÈME 5. *Une fonction continue sur un intervalle $[a, b]$ et ayant une dérivée positive (resp. strictement positive) sur $]a, b[$ est croissante (resp. strictement croissante) sur $[a, b]$.*

[1]) Louis de Lagrange, mathématicien français (1736-1813).

Soit en effet $a \leqslant x_1 < x_2 \leqslant b$. Les hypothèses du théorème de Lagrange étant remplies sur $[x_1, x_2]$, il existe un point $c \in ]x_1, x_2[$, tel que

$$f(x_2) - f(x_1) = (x_2 - x_1) f'(c).$$

Si $f' \geqslant 0$ sur $]a, b[$, alors $f'(c) \geqslant 0$ et

$$f(x_2) - f(x_1) \geqslant 0; \tag{7}$$

si $f' > 0$ sur $]a, b[$, alors $f'(c) > 0$ et

$$f(x_2) - f(x_1) > 0. \tag{8}$$

Comme les inégalités (7) et (8) ont lieu pour tous $x_1$ et $x_2$, tels que $a \leqslant x_1 < x_2 \leqslant b$, la fonction $f$ est croissante sur $[a, b]$ dans le premier cas et strictement croissante dans le second.

EXEMPLE 1. Revenons à l'exemple du § 7, où il fallait évaluer la quantité $\lambda = \sqrt[3]{8{,}001} - \sqrt[3]{8}$. En appliquant la formule des accroissements finis à la fonction $\psi(x) = x^{1/3}$, on obtient

$$\lambda = \psi(8{,}001) - \psi(8) = 0{,}001 \cdot \psi'(c) = 0{,}001 \cdot \frac{1}{3} x^{-2/3}|_{x=c} =$$

$$= \frac{1}{3000} c^{-2/3} < \frac{1}{3000} 8^{-2/3} = \frac{1}{12\,000}.$$

Le résultat est le même, mais ici il est entièrement justifié.

EXEMPLE 2. La fonction $y = \operatorname{sh} x = \frac{1}{2}(e^x - e^{-x})$ a une dérivée continue

$$(\operatorname{sh} x)' = \frac{1}{2}(e^x + e^{-x}) = \operatorname{ch} x > 0, \quad \forall x \in ]-\infty, \infty[,$$

et est telle que

$$\lim_{x \to -\infty} \operatorname{sh} x = -\infty, \quad \lim_{x \to +\infty} \operatorname{sh} x = +\infty.$$

Donc, elle est strictement croissante, continûment dérivable sur $]-\infty, \infty[$ et applique $]-\infty, \infty[$ sur lui-même. Elle admet par suite une fonction réciproque continûment dérivable:

$$x = \operatorname{Arg\,sh} y, \quad y \in ]-\infty, \infty[.$$

THÉORÈME 6. *Si une fonction admet une dérivée nulle sur un intervalle* $]a, b[$, *elle est constante sur cet intervalle.*

En effet, en vertu du théorème de Lagrange, on a

$$f(x) - f(x_1) = (x - x_1) f'(c),$$

où $x_1$ est un point fixe de l'intervalle $]a, b[$, $x$ un point quelconque de $]a, b[$ (situé à gauche ou à droite de $x_1$) et $c$ un point compris entre $x_1$ et $x$ et dépendant d'eux. Comme $f'(x) \equiv 0$ sur $]a, b[$ par hypothèse, on a $f'(c) = 0$ et $f(x) = f(x_1) = C$ pour tous les $x \in ]a, b[$.

Signalons que si les hypothèses des théorèmes ci-dessus sont affaiblies, ces derniers peuvent être mis en défaut (voir remarques 1, 2 suivant le théorème de Rolle).

DÉFINITION. *On dira qu'une fonction $y = f(x)$ est strictement croissante (resp. strictement décroissante) en un point $x_0$ s'il existe un nombre $\delta > 0$ tel que*

$$\frac{\Delta y}{\Delta x} > 0 \left(resp. \frac{\Delta y}{\Delta x} < 0\right), \quad 0 < |\Delta x| < \delta.$$

THÉORÈME 7. *Si $f'(x_0) > 0$ (resp. $< 0$), la fonction $f(x)$ est strictement croissante (resp. strictement décroissante) en $x_0$.*

DÉMONSTRATION. Comme $f'(x_0) = \lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$, alors pour $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que $f'(x_0) - \varepsilon < \frac{\Delta y}{\Delta x} < f'(x_0) + \varepsilon$, pourvu que $|\Delta x| < \delta$. Supposons que $f'(x_0) > 0$. En prenant $\varepsilon < f'(x_0)$, on obtient $\frac{\Delta y}{\Delta x} > 0$ pour $|\Delta x| < \delta$, c'est-à-dire que la fonction $f$ est strictement croissante en $x_0$.

REMARQUE 6. Si une fonction $f$ a une dérivée et croît sur $]a, b[$, alors $f'(x) \geqslant 0$ sur $]a, b[$. En effet, il est impossible dans les conditions indiquées que la dérivée de $f$ soit négative en un point $x \in ]a, b[$, car cela contredirait le théorème 7.

Si la seule information dont on dispose sur une fonction $f$ est qu'elle a une dérivée et croît strictement sur un intervalle $]a, b[$, on doit tout de même conclure que $f'(x) \geqslant 0$ sur $]a, b[$, car une fonction strictement croissante peut éventuellement avoir une dérivée nulle en certains points de $]a, b[$. C'est le cas notamment de la fonction $x^3$ qui est strictement croissante sur $]-\infty, \infty[$, et dont la dérivée est nulle en $x = 0$.

REMARQUE 7. Si une fonction est strictement croissante en un point $x_0$, elle ne l'est pas automatiquement en un voisinage de $x_0$.

Citons pour exemple la fonction

$$F(x) = \begin{cases} 0, & x = 0, \\ \frac{x}{2} - x^2 \sin \frac{1}{x}, & x \neq 0. \end{cases}$$

Il est évident que

$$F'(0) = \lim_{x \to 0} \frac{\frac{x}{2} - x^2 \sin \frac{1}{x}}{x} = \frac{1}{2}$$

et $F(x)$ est strictement croissante en $x = 0$. Cette fonction n'est pourtant pas monotone, car la dérivée $F'(x) = \frac{1}{2} - 2x \sin \frac{1}{x} + \cos \frac{1}{x}$ est tantôt positive, tantôt négative en tout voisinage de 0 (voir théorème 5). Pour $x_k = 1/k\pi$ ($k = 1, 2, \ldots$) elle est égale à $3/2$ si $k$ est pair et à $-1/2$ s'il est impair.

THÉORÈME 8. *Si une fonction $f(x)$ est paire (resp. impaire) et dérivable sur $[-a, a]$, alors sa dérivée $f'(x)$ est impaire (resp. paire).*

DÉMONSTRATION. Comme $f(x) \equiv f(-x)$ sur $[-a, a]$, on a $f'(x) \equiv -f'(-x)$, ce qui exprime que la dérivée $f'(x)$ est impaire. (On démontre ce fait en se servant aussi de la définition de la dérivée.)

## § 13. Levée des indéterminations

On dira que l'expression $\frac{f(x)}{g(x)}$ donne lieu à une indétermination de la forme $\frac{0}{0}$ pour $x \to a$, si $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$. Lever cette indétermination, c'est trouver $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ si elle existe.

THÉORÈME 1. *Supposons que $f(x)$ et $g(x)$ sont définies et dérivables au voisinage d'un point $x = a$, sauf éventuellement en $a$, et que*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0, \quad g(x) \neq 0 \quad et \quad g'(x) \neq 0$$

*dans ce voisinage. Sous ces conditions, l'existence de $\lim\limits_{x\to a} \frac{f'(x)}{g'(x)}$ entraîne celle de $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ et*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}. \tag{1}$$

DÉMONSTRATION. On admettra que $a$ est un nombre fini. (Si $a = \infty$, voir remarque 3 plus bas.) Prolongeons les fonctions $f$ et $g$ au point $x = a$ en posant $f(a) = g(a) = 0$. Ces fonctions seront continues en $a$. Considérons l'intervalle $[a, x]$, où $x > a$ ou $x < a$ (voir remarque 5 du § 12). Les fonctions $f$ et $g$ sont continues sur $[a, x]$ et dérivables sur $]a, x[$, donc en vertu du théorème de Cauchy il existe un point $\xi$ tel que

$$\frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(\xi)}{g'(\xi)} \ (\xi \in ]a, x[) \quad \text{ou} \quad \frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}.$$

Si $x \to a$, il en est de même de $\xi$, et

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{\xi\to a} \frac{f'(\xi)}{g'(\xi)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} \tag{2}$$

sous réserve que ces limites existent. C.q.f.d.

REMARQUE 1. $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ peut exister même si $\lim\limits_{x\to a} \frac{f'(x)}{g'(x)}$ n'existe pas.

EXEMPLE 1. Comme $\sin x \approx x$, il vient

$$\lim_{x\to 0} \frac{x^2 \sin \frac{1}{x}}{\sin x} = \lim_{x\to 0} x \sin \frac{1}{x} = 0,$$

tandis que

$$\lim_{x\to 0}\frac{\left(x^2\sin\frac{1}{x}\right)'}{(\sin x)'}=\lim_{x\to 0}\frac{2x\sin\frac{1}{x}-\cos\frac{1}{x}}{\cos x}$$

n'existe pas.

Remarque 2. Si l'expression $\frac{f'(x)}{g'(x)}$ donne lieu à une indétermination de la forme $\frac{0}{0}$ et que les fonctions $f'(x)$ et $g'(x)$ satisfassent aux hypothèses du théorème 1, alors

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}=\lim_{x\to a}\frac{f''(x)}{g''(x)}.$$

Ces égalités expriment que l'existence de la troisième limite entraîne celle de la seconde et de la première.

Théorème 2 $\left(\frac{\infty}{\infty}\right)$. *Supposons que $f$ et $g$ sont définies et dérivables au voisinage d'un point $x=a$, que $\lim\limits_{x\to a}f(x)=\lim\limits_{x\to a}g(x)=\infty$ et que $g(x)\neq 0$ et $g'(x)\neq 0$ dans ce voisinage. Alors l'existence de $\lim\limits_{x\to a}\frac{f'(x)}{g'(x)}$ entraîne celle de $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$ et*

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}.$$

Nous glisserons sur la démonstration de ce théorème.

Remarque 3. Si $a=\infty$, la substitution $x=1/t$ nous ramène au cas $a=0$:

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\lim_{t\to 0}\frac{f(1/t)}{g(1/t)}=\lim_{t\to 0}\frac{(f(1/t))'}{(g(1/t))'}=\lim_{t\to 0}\frac{f'(1/t)(-1/t^2)}{g'(1/t)(-1/t^2)}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)}.$$

La règle exprimée par les théorèmes 1 et 2 et qui dit que la limite du rapport de deux fonctions est égale à la limite du rapport de leurs dérivées s'appelle *règle de l'Hospital* du nom du mathématicien qui la formula le premier pour des cas élémentaires. Du reste cette règle était connue de Bernoulli avant l'Hospital [1]).

Exemple 2.

$$\lim_{x\to\infty}\frac{x^\alpha}{a^x}=0,\quad \forall\alpha>0,\quad a>1.$$

[1]) Guillaume de l'Hospital, mathématicien français (1661-1704), Jean Bernoulli, mathématicien suisse (1667-1748).

Nous avons une forme indéterminée de type $\left(\frac{\infty}{\infty}\right)$. En appliquant $k$ fois la règle de l'Hospital ($k \geqslant \alpha$ ; si $\alpha$ est entier, $k = \alpha$), on obtient

$$\lim_{x\to\infty} \frac{x^\alpha}{a^x} = \lim_{x\to\infty} \frac{cx^{\alpha-k}}{a^x (\ln a)^k} = 0.$$

EXEMPLE 3.

$$\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = 0, \quad \forall\alpha > 0.$$

Les fonctions $x^\alpha$ et $\ln x$ satisfont à toutes les conditions du théorème 2, donc

$$\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = \lim_{x\to\infty} \frac{1/x}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha x^\alpha} = 0.$$

On rencontre aussi des formes indéterminées de type $0\cdot\infty$, $0^0$, $\infty^0$, $\infty - \infty$, $1^\infty$. Ces formes se ramènent à celles de type $\frac{0}{0}$ ou $\frac{\infty}{\infty}$ par des transformations algébriques.

a. *Forme indéterminée de type* $0\cdot\infty$ ($f(x)\, g(x)$, $f(x) \to 0$, $g(x) \to \infty$ pour $x \to a$). Il est clair que

$$f(x)\, g(x) = \frac{f}{1/g} \left(\frac{0}{0}\right) \quad \text{ou} \quad f\cdot g = \frac{g}{1/f} \left(\frac{\infty}{\infty}\right).$$

EXEMPLE 4.

$$\lim_{x\to 0} x^\alpha \ln x = 0, \quad \forall\alpha > 0\,;$$

$$\lim_{x\to 0} x^\alpha \ln x = \lim_{x\to 0} \frac{\ln x}{x^{-\alpha}} = \left(\frac{\infty}{\infty}\right) = \lim_{x\to 0} \frac{1/x}{-\alpha x^{-\alpha-1}} = -\frac{1}{\alpha} \lim_{x\to 0} x^\alpha = 0.$$

b. Les *formes indéterminées de type* $1^\infty$, $0^0$, $\infty^0$ se ramènent à celle de type $0\cdot\infty$. En effet, $f^g = e^{g\ln f}$ ($f > 0$).

Si

$$\lim_{x\to a} g \ln f = k,$$

alors

$$\lim_{x\to a} f^g = e^k.$$

c. *Forme indéterminée de type* $\infty - \infty$ ($f(x) - g(x)$, $f \to +\infty$, $g \to +\infty$ pour $x \to a$). Il est immédiat de voir que

$$f - g = \frac{1}{\frac{1}{f}} - \frac{1}{\frac{1}{g}} = \frac{\frac{1}{g} - \frac{1}{f}}{\frac{1}{f}\cdot\frac{1}{g}} \left(\frac{0}{0}\right).$$

## § 14. Formule de Taylor [1])

Soit un polynôme de degré $n$:

$$P_n(x) = b_0 + b_1 x + \ldots + b_n x^n = \sum_{k=0}^{n} b_k x^k.$$

En posant $x = (x - x_0) + x_0$, où $x_0$ est un nombre fixe, on obtient

$$P_n(x) = \sum_{k=0}^{n} b_k [(x - x_0) + x_0]^k. \qquad (1)$$

En réduisant les termes semblables par rapport aux puissances de $x - x_0$, on obtient l'expression suivante:

$$P_n(x) = a_0 + a_1(x - x_0) + \ldots + a_n(x - x_0)^n = \sum_{k=0}^{n} a_k (x - x_0)^k, \qquad (2)$$

appelée *développement du polynôme $P_n(x)$ en série entière de $x - x_0$* ou encore *au voisinage de $x_0$*. Les coefficients $a_0, a_1, \ldots, a_n$ dépendent de $b_i$ et de $x_0$. Par exemple, $a_0 = b_0 + b_1 x_0 + \ldots + b_n x_0^n$. Sur (1) on voit de toute évidence que $P_n(x)$ ne dépend pas de $x_0$.

Calculons les dérivées successives de $P_n(x)$:

$$\left\{\begin{array}{l} P'_n(x) = a_1 + 2a_2(x - x_0) + \ldots + na_n(x - x_0)^{n-1}, \\ P''_n(x) = 1\cdot 2a_2 + 2\cdot 3a_3(x - x_0) + \ldots + n(n-1)a_n(x - x_0)^{n-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ P_n^{(k)}(x) = 1\cdot 2\ldots ka_k + \ldots + n(n-1)\ldots(n-k+1)\times \\ \qquad\qquad\qquad\qquad\qquad\qquad \times a_n(x - x_0)^{n-k}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ P_n^{(n)}(x) = 1\cdot 2\ldots na_n. \end{array}\right. \qquad (3)$$

Les dérivées d'ordre supérieur à $n$ sont nulles. En faisant $x = x_0$ dans les formules (2) et (3), on obtient

$$P_n(x_0) = a_0, \quad P'_n(x_0) = a_1,$$

$$P''_n(x_0) = 1\cdot 2a_2, \ \ldots, \ P_n^{(k)}(x_0) = k!\, a_k, \ \ldots, \ P_n^{(n)}(x_0) = n!a_n$$

ou

$$a_k = \frac{P_n^{(k)}(x_0)}{k!} \quad (k = 0, 1, \ldots, n), \qquad (4)$$

où l'on convient que $0! = 1$, $P_n^{(0)}(x) = P_n(x)$.

Les formules (4) montrent qu'un polynôme $P_n(x)$ admet un développement unique suivant les puissances de $x - x_0$, c'est-à-dire que si

$$P_n(x) = \sum_{k=0}^{n} \beta_k (x - x_0)^k = \sum_{k=0}^{n} \beta'_k (x - x_0)^k,$$

[1]) Brook Taylor, mathématicien anglais (1685-1731).

où $\beta_k$ et $\beta'_k$ sont des constantes, alors $\beta_k = \beta'_k$ $(k = 0, 1, \ldots, n)$. En effet, les nombres $\beta_k$ et $\beta'_k$ se calculent avec la même formule (4).

Compte tenu de (4), la formule (2) peut encore s'écrire:

$$P_n(x) = P_n(x_0) + \frac{P'_n(x_0)}{1!}(x - x_0) + \ldots$$

$$\ldots + \frac{P_n^{(n)}(x_0)}{n!}(x - x_0)^n = \sum_{k=0}^{n} \frac{P_n^{(k)}(x_0)}{k!}(x - x_0)^k. \quad (2')$$

La formule (2') s'appelle *formule de Taylor pour le polynôme* $P_n(x)$. On remarquera que le second membre de (2') ne dépend pas de $x_0$.

Exemple 1. Soient $P_n(x) = (a + x)^n$ et $x_0 = 0$. En vertu de (2'), on a

$$P_n(x) = \sum_{k=0}^{n} \frac{P_n^{(k)}(0)}{k!} x^k,$$

où

$$P_n^{(k)}(x) = n(n-1) \ldots (n-k+1)(a+x)^{n-k},$$
$$P_n^{(k)}(0) = n(n-1) \ldots (n-k+1)\, a^{n-k},$$

et l'on obtient la formule du *binôme de Newton*

$$(a+x)^n = \sum_{k=0}^{n} \frac{n(n-1)\ldots(n-k+1)}{k!} a^{n-k} x^k. \quad (5)$$

Considérons maintenant une fonction quelconque $f(x)$ possédant des dérivées jusqu'à l'ordre $n + 1$ compris continues dans un voisinage d'un point $x_0$. L'expression

$$Q_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \quad (6)$$

s'appelle *développement de Taylor d'ordre n de la fonction f en série entière de* $x - x_0$ ou *au voisinage de* $x_0$.

Le polynôme $Q_n(x)$ est confondu avec la fonction $f(x)$ au point $x_0$ seulement; pour tous les autres $x$ on a $Q_n(x) \neq f(x)$ (sauf si $f(x)$ est un polynôme de degré $n$). De plus

$$Q'_n(x_0) = f'(x_0), \ldots, Q_n^{(n)}(x_0) = f^{(n)}(x_0). \quad (7)$$

Posons

$$f(x) = Q_n(x) + r_n(x). \quad (8)$$

La formule (8) s'appelle *formule de Taylor pour la fonction* $f(x)$; $r_n(x)$ est le *reste de la formule de Taylor*. La fonction $r_n(x)$ mesure l'erreur commise en remplaçant $f(x)$ par le polynôme (6).

Exprimons $r_n(x)$ en fonction de la dérivée $f^{(n+1)}(x)$.

En vertu de (7) et de (8), on a $r_n(x_0) = r'_n(x_0) = \ldots = r_n^{(n)}(x_0) = 0$. Posons $\varphi(x) = (x - x_0)^{n+1}$. Il est évident que $\varphi(x_0) = \varphi'(x_0) = \ldots = \varphi^{(n)}(x_0) = 0$. En appliquant le théorème de Cauchy aux fonctions $r_n(x)$ et $\varphi(x)$, on aura

$$\frac{r_n(x)}{\varphi(x)} = \frac{r_n(x) - r_n(x_0)}{\varphi(x) - \varphi(x_0)} = \frac{r'_n(x_1)}{\varphi'(x_1)} = \frac{r'_n(x_1) - r'_n(x_0)}{\varphi'(x_1) - \varphi'(x_0)} = \frac{r''_n(x_2)}{\varphi''(x_2)} = \ldots$$

$$\ldots = \frac{r_n^{(n)}(x_n)}{\varphi^{(n)}(x_n)} = \frac{r_n^{(n)}(x_n) - r_n^{(n)}(x_0)}{\varphi^{(n)}(x_n) - \varphi^{(n)}(x_0)} = \frac{r_n^{(n+1)}(x_{n+1})}{\varphi^{(n+1)}(x_{n+1})}$$

($x_1 \in ]x_0, x[$ et $x_{k+1} \in ]x_0, x_k[$, $k = 1, 2, \ldots, n$).

Or

$$\varphi^{(n+1)}(x) = (n+1)!, \quad r_n^{(n+1)}(x) = f^{(n+1)}(x) - 0 = f^{(n+1)}(x).$$

Donc

$$r_n(x) = \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(c), \tag{9}$$

où $c = x_{n+1}$ est un point compris entre $x_0$ et $x$.

La formule (8) peut encore s'écrire

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}. \tag{8'}$$

Dans la formule (8') le reste est mis sous la *forme de Lagrange*.

Nous avons prouvé un théorème important.

THÉORÈME 1. *Si une fonction $f$ admet une dérivée $f^{(n+1)}(x)$ continue au voisinage d'un point $x_0$, alors pour tout point $x$ de ce voisinage il existe un point $c \in ]x_0, x[$ tel que $f(x)$ se représente par la formule* (8').

Le point $c$ dépend de $x$ et de $n$.

Si $x_0 = 0$, la formule (8) s'appelle *formule de Maclaurin*.

Le reste de la formule de Taylor se représente sous plusieurs formes. Signalons l'importante *forme de Cauchy*

$$r_n(x) = \frac{(x - x_0)^{n+1}(1 - \theta)^n}{n!} f^{(n+1)}(x_0 + \theta(x - x_0)), \tag{10}$$

où $\theta \in ]0, 1[$ dépend de $n$ et de $x$. Cette formule sera établie au § 5 du chap. 6.

En réduisant le voisinage du point $x_0$, on obtient un intervalle *fermé* $[x_0 - \delta, x_0 + \delta]$ sur lequel est continue la dérivée $f^{(n+1)}(x)$. Elle est donc bornée sur cet intervalle par un nombre strictement positif $M_n$ dépendant de $n$ mais pas de $x$:

$$|f^{(n+1)}(x)| \leqslant M_n, \quad x_0 - \delta \leqslant x \leqslant x_0 + \delta \tag{11}$$

(voir chap. 3, § 5, théorème 1). Alors

$$|r_n(x)| \leqslant \frac{|f^{(n+1)}(c)|}{(n+1)!}|x-x_0|^{n+1} \leqslant$$
$$\leqslant \frac{M_n|x-x_0|^{n+1}}{(n+1)!}, \quad |x-x_0| < \delta. \tag{12}$$

L'inégalité (12) peut servir à deux fins: d'une part à étudier le comportement de $r_n(x)$ pour $n$ fixe au voisinage de $x_0$, de l'autre à étudier le comportement de $r_n(x)$ pour $n \to \infty$.

De (12) il s'ensuit par exemple que pour $n$ fixe

$$r_n(x) = o((x-x_0)^n), \quad x \to x_0, \tag{13}$$

relation qui indique que le quotient de $r_n(x)$ par $(x-x_0)^n$ converge vers 0 pour $x \to x_0$.

De (8′) et compte tenu de (13), il vient

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \underset{x\to x_0}{o}((x-x_0)^n). \tag{14}$$

Cette formule s'appelle *formule de Taylor avec un reste de Peano* [1]). Elle est favorable à l'étude de la fonction $f$ au voisinage du point $x_0$.

THÉORÈME 2 (D'UNICITÉ). *Si une fonction $f$ se représente au voisinage d'un point $x_0$ par les formules*

$$\begin{cases} f(x) = a_0 + a_1(x-x_0) + \ldots + a_n(x-x_0)^n + \underset{x\to x_0}{o}((x-x_0)^n), \\ f(x) = b_0 + b_1(x-x_0) + \ldots + b_n(x-x_0)^n + \underset{x\to x_0}{o}((x-x_0)^n), \end{cases} \tag{15}$$

*alors*

$$a_k = b_k \quad (k = 0, 1, \ldots, n). \tag{16}$$

DÉMONSTRATION. En égalant les seconds membres de (15) et en passant à la limite pour $x \to x_0$, on obtient $a_0 = b_0$. Si maintenant on divise les deux membres de l'égalité obtenue par $x - x_0$ ($x \neq x_0$) et que l'on passe ensuite à la limite pour $x \to x_0$, on obtient $a_1 = b_1$. En poursuivant cette procédure, on obtient en définitive $a_n = b_n$.

EXEMPLE 2. On sait que

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} \quad (x \neq 1).$$

Donc

$$\psi(x) = \frac{1}{1-x} = \sum_{k=0}^{n} x^k + \frac{x^{n+1}}{1-x} = \sum_{k=0}^{n} x^k + \underset{x\to 0}{o}(x^n). \tag{17}$$

[1]) Giuseppe Peano, mathématicien italien (1858-1932).

La fonction $\psi$ ayant des dérivées de tout ordre au voisinage du point $x = 0$, elle est justiciable de la formule de Taylor avec un reste de Peano

$$\psi(x) = \sum_{k=0}^{n} \frac{\psi^{(k)}(0)}{k!} x^k + \underset{x\to 0}{o}(x^n). \tag{18}$$

En comparant les formules (17) et (18), on obtient en vertu du théorème d'unicité que

$$1 = \frac{\psi^{(k)}(0)}{k!} \quad (k = 0, 1, \ldots, n). \tag{19}$$

Dans le paragraphe suivant on étudie le comportement du reste de la formule de Taylor pour $n \to \infty$.

## § 15. Série de Taylor

On appelle *série* une expression de la forme

$$a_0 + a_1 + \ldots \tag{1}$$

ou encore

$$\sum_{k=0}^{\infty} a_k, \tag{1'}$$

où $a_k$ sont des nombres dépendant de l'indice $k$. Les sommes finies

$$S_n = \sum_{k=0}^{n} a_k \quad (n = 0, 1, 2, \ldots)$$

s'appellent *sommes partielles* de la série (1) (ou (1')). Si existe la limite finie

$$\lim_{n\to\infty} S_n = S, \tag{2}$$

on dit que la série (1) converge vers le nombre $S$, et $S$ s'appelle *somme* de la série. On note

$$S = \sum_{k=0}^{\infty} a_k = a_0 + a_1 + a_2 + \ldots$$

On dit que la série (1) est *divergente* si la limite des sommes partielles $S_n$ $(n \to \infty)$ n'existe pas ou est égale à $\infty$.

Supposons maintenant que la fonction $f$ a des dérivées de tout ordre au voisinage d'un point $x_0$. La série

$$f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots \tag{3}$$

ou de façon plus concise

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \tag{3'}$$

s'appelle *série de Taylor* de $x - x_0$ associée à la fonction $f$. Si $x_0 = 0$. cette série s'appelle *série de Maclaurin*.

Un cas particulièrement intéressant est celui où la série de Taylor de $x - x_0$ de la fonction $f$ converge au voisinage de $x_0$ vers la fonction $f(x)$. Si ceci a lieu, alors

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \quad x \in ]x_0 - \delta, \ x_0 + \delta[,$$

c'est-à-dire que la fonction $f(x)$ est la somme de sa série de Taylor au voisinage de $x_0$. On dit alors que *la fonction $f(x)$ se développe au voisinage de $x_0$ en une série de Taylor convergeant vers $f(x)$.*

Théorème 1. *Si une fonction $f$ possède des dérivées de tout ordre sur un intervalle $[x_0 - \delta, \ x_0 + \delta]$ et si le reste de sa formule de Taylor tend vers 0 pour $n \to \infty$, i.e.*

$$\lim_{n \to \infty} r_n(x) = 0, \ x \in [x_0 - \delta, \ x_0 + \delta], \tag{4}$$

*alors $f$ se développe en une série de Taylor convergeant vers elle sur cet intervalle.*

Démonstration. Supposons que $f$ possède des dérivées de tout ordre sur $[x_0 - \delta, \ x_0 + \delta]$. Ces dérivées sont continues sur $[x_0 - \delta, \ x_0 + \delta]$, car si $f$ possède une dérivée $f^{(k)}$ sur $[x_0 - \delta, \ x_0 + \delta]$, alors $f^{(k-1)}$ est continue sur $[x_0 - \delta, \ x_0 + \delta]$.

Donc, la fonction $f$ est justiciable de la formule de Taylor

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + r_n(x), \quad \forall n, \ x \in [x_0 - \delta, \ x_0 + \delta].$$

En vertu de (4)

$$\lim_{n \to \infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \lim_{n \to \infty} [f(x) - r_n(x)] =$$

$$= f(x) - \lim_{n \to \infty} r_n(x) = f(x),$$

c'est-à-dire que le développement de Taylor de la fonction $f(x)$ converge pour $n \to \infty$ vers $f(x)$:

$$\lim_{n \to \infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = f(x), \quad x \in [x_0 - \delta, \ x_0 + \delta]. \tag{5}$$

Or, cela signifie que la série de Taylor de la fonction $f(x)$ converge sur $[x_0 - \delta, \ x_0 + \delta]$ et sa somme est $f(x)$:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \quad x \in [x_0 - \delta, \ x_0 + \delta].$$

Ce que nous voulions.

Le théorème suivant exprime une condition suffisante simple de convergence du reste de la formule de Taylor vers 0.

THÉORÈME 2. *Si une fonction f possède sur un intervalle* $[x_0 - \delta, x_0 + \delta]$ *des dérivées de tout ordre, bornées par le même nombre M* ($|f^{(n)}(x)| \leqslant M$, $n = 0, 1, 2, \ldots$), *alors le reste de sa formule de Taylor converge vers 0 pour* $n \to \infty$ *sur cet intervalle*:

$$\lim_{n\to\infty} r_n(x) = 0. \tag{6}$$

DÉMONSTRATION. On se servira du reste de Lagrange. On a

$$|r_n(x)| = \frac{|x - x_0|^{n+1}}{(n+1)!}\,|f^{(n+1)}(c)| \leqslant \frac{M\cdot\delta^{n+1}}{(n+1)!}, \tag{7}$$

où $c \in ]x_0, x[$, $M \geqslant |f^{(n+1)}(x)|$, $\forall n \in \mathbb{N}$ et $|x - x_0| < \delta$.

Comme le second membre de (7) tend vers 0 pour $n \to \infty$ (voir chap. 2, § 5, (5)), on obtient (6).

## § 16. Formules et séries de Taylor des fonctions élémentaires

1. $f(x) = e^x$. Cette fonction est indéfiniment dérivable sur $]-\infty, \infty[$. De plus

$$f^{(k)}(x) = e^x, \quad f^{(k)}(0) = 1 \quad (k = 0, 1, \ldots), \quad f^{(n+1)}(c) = e^c.$$

La formule de Taylor avec un reste de Lagrange s'écrit

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + r_n(x), \quad r_n(x) = \frac{e^c x^{n+1}}{(n+1)!}, \quad c \in ]0, x[, \tag{1}$$

où $x$ est positif ou négatif. Sur l'intervalle $[-A, A]$, $A > 0$, on a

$$|r_n(x)| \leqslant \frac{e^A A^{n+1}}{(n+1)!} \to 0, \quad n \to \infty. \tag{2}$$

Ceci montre (voir théorème 1 du § 15) que la fonction $e^x$ se développe sur $[-A, A]$ en une série de Taylor de $x$ (ou série de Maclaurin) convergeant vers $e^x$:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{3}$$

Or, $A > 0$ est un nombre arbitraire, donc cette relation est valable sur l'axe numérique tout entier ($x \in ]-\infty, \infty[$). Ici $|f^{(k)}(x)| = |e^x| \leqslant e^A$ ($k = 0, 1, 2, \ldots$) sur $[-A, A]$, et pour déduire (3) on aurait pu faire intervenir le théorème 2 du § 15.

Calculons le nombre $e$ à 0,001 près. On a (voir (1))

$$e = \sum_{k=0}^{n} \frac{1}{k!} + r_n(1), \tag{4}$$

où

$$r_n(1)=\frac{e^c}{(n+1)!}, \quad 0<c<1. \tag{5}$$

Il faut choisir $n$ assez grand pour que

$$r_n(1)=\frac{e^c}{(n+1)!}\leqslant 0{,}001 \quad (0<c<1).$$

Comme $e^c<3$, il suffit seulement de résoudre l'inégalité $3/(n+1)!\leqslant \leqslant 0{,}001$. Celle-ci est réalisée pour $n=6$. Donc,

$$e\approx 2+\frac{1}{2!}+\frac{1}{3!}+\ldots+\frac{1}{6!}=2{,}718$$

à 0,001 près.

REMARQUE. Comme $1<e^c<3$ pour $0<c<1$, alors pour $n>2$ on a $e^c/(n+1)=\theta$, où $0<\theta<1$. On peut donc mettre l'égalité (4) sous la forme

$$e=\sum_{k=0}^{n}\frac{1}{k!}+\frac{\theta}{n!}.$$

On s'est servi de cette formule (formule (3), § 6, chap. 2) pour prouver que $e$ est irrationnel.

2. $y=\sin x$. Cette fonction est indéfiniment dérivable et

$$|(\sin x)^{(k)}|=\left|\sin\left(x+k\frac{\pi}{2}\right)\right|\leqslant 1, \quad \forall k\in\mathbf{N}.$$

Donc, la fonction $\sin x$ se développe d'après le théorème 2 du paragraphe précédent en une série de Taylor de $x$ convergeant vers elle sur $]-\infty, \infty[$:

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots=\sum_{k=0}^{\infty}\frac{(-1)^k x^{2k+1}}{(2k+1)!}.$$

On tiendra compte du fait que

$$(\sin x)^{(n)}|_{x=0}=\sin\frac{\pi n}{2}=\begin{cases}0 & \text{pour } n=2k,\\ (-1)^k & \text{pour } n=2k+1.\end{cases}$$

La formule de Taylor de $\sin x$ s'écrit

$$\sin x=x-\frac{x^3}{3!}+\ldots+(-1)^{\nu+1}\frac{x^{2\nu-1}}{(2\nu-1)!}+r_{2\nu}(x), \tag{6}$$

où

$$r_{2\nu}(x)=\frac{x^{2\nu+1}}{(2\nu+1)!}\sin\left(\theta x+(2\nu+1)\frac{\pi}{2}\right), \quad 0<\theta<1.$$

On remarque que

$$r_{2\nu}(x)=\underset{x\to 0}{o}(x^{2\nu})$$

et par suite

$$\sin x = x - \frac{x^3}{3!} + \ldots + (-1)^{\nu+1} \frac{x^{2\nu-1}}{(2\nu-1)!} + \underset{x \to 0}{o}(x^{2\nu}).$$

3. $y = \cos x$. On obtient par analogie

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots = \sum_{k=0}^{\infty} (-1^k) \frac{x^{2k}}{(2k)!}.$$

EXEMPLE 1. Trouver

$$\lim_{x \to 0} \frac{\sin x - x}{x^3}.$$

On a

$$\sin x = x - \frac{x^3}{3!} + o(x^3), \tag{7}$$

donc

$$\frac{\sin x - x}{x^3} = -\frac{1}{3!} + \frac{\underset{x \to 0}{o}(x^3)}{x^3} = -\frac{1}{3!} + \underset{x \to 0}{o}(1) \to -\frac{1}{3!},$$

c'est-à-dire que

$$\lim_{x \to 0} \frac{\sin x - x}{x^3} = -\frac{1}{6}.$$

En fait, le reste est de la forme $\underset{x \to 0}{o}(x^4)$ dans (7). Or, il suffit qu'il soit de la forme $\underset{x \to 0}{o}(x^3)$. Il est classique que $\underset{x \to 0}{o}(x^4)$ entraîne automatiquement $\underset{x \to 0}{o}(x^3)$ (la réciproque n'est pas vraie).

4. $f(x) = \ln(1 + x)$. Cette fonction est définie et indéfiniment dérivable pour $x > -1$. Donc, elle est justiciable de la formule de Taylor pour tout $n = 1, 2, \ldots$ Comme

$$f^{(n)}(x) = \frac{(-1)^{n+1}(n-1)!}{(1+x)^n}, \quad f^{(n)}(0) = (-1)^{n+1}(n-1)!,$$

la formule de Taylor s'écrit

$$\ln(1+x) = x - \frac{x^2}{2} + \ldots + (-1)^{n+1} \frac{x^n}{n} + r_n(x).$$

En se servant des formes de Lagrange et de Cauchy du reste on démontre que

$$\lim_{n \to \infty} r_n(x) = 0 \text{ pour } -1 < x \leqslant 1.$$

Donc, la série de Taylor de la fonction $\ln(1 + x)$ est

$$\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} \quad (-1 < x \leqslant 1).$$

5. $f(x) = (1+x)^m$. On a

$$f^{(n)}(x) = m(m-1)\ldots(m-n+1)(1+x)^{m-n},$$
$$f^{(n)}(0) = m(m-1)\ldots(m-n+1).$$

La série de Taylor s'écrit

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \ldots$$
$$\ldots + \frac{m(m-1)\ldots(m-n+1)}{n!}x^n + r_n(x).$$

On démontre que pour tout $m$

$$\lim_{n\to\infty} r_n(x) = 0 \quad (-1 < x < 1).$$

Donc, pour tout $m$ réel, la série de Taylor de la fonction $(1+x)^m$ est

$$(1+x)^m = 1 + \sum_{k=1}^{\infty} \frac{m(m-1)\ldots(m-k+1)}{k!}x^k \quad (-1 < x < 1). \tag{8}$$

Si $m$ est un entier naturel, la fonction $(1+x)^m$ est un polynôme. Dans ce cas $r_n(x) \equiv 0$ pour $n > m$ et la série de droite de (8) est un développement limité de Taylor (voir § 14).

EXEMPLE 2. Calculer la limite ($m \neq n$, $m \neq 0$, $n \neq 0$)

$$\lim_{x\to 0} \frac{\sqrt[m]{1+x} - \sqrt[n]{1+x}}{x} = \lim_{x\to 0} \frac{1 + \frac{x}{m} + o(x) - \left(1 + \frac{x}{n} + o(x)\right)}{x} =$$
$$= \lim_{x\to 0} \frac{x\left(\frac{1}{m} - \frac{1}{n}\right) + o(x)}{x} = \lim_{x\to 0}\left[\left(\frac{1}{m} - \frac{1}{n}\right) + o(1)\right] = \frac{1}{m} - \frac{1}{n}.$$

EXEMPLE 3.

$$\lim_{x\to 0} \frac{\ln(1+x) - x(1+x)^{\alpha}}{x^2} = \lim_{x\to 0} \frac{x - \frac{x^2}{2} + o(x^2) - x(1+\alpha x + o(x))}{x^2} =$$
$$= \lim_{x\to 0} \frac{-\left(\frac{1}{2} + \alpha\right)x^2 + o(x^2)}{x^2} = \lim_{x\to 0}\left[-\frac{1}{2} - \alpha + o(1)\right] = -\frac{1}{2} - \alpha.$$

## § 17. Extrémum local d'une fonction

On a déjà défini l'extrémum local au début du paragraphe 12. On peut énoncer cette définition sous la forme :

*On dit qu'une fonction $y = f(x)$ a un maximum (resp. minimum) local en un point $c$ si l'on peut exhiber un $\delta > 0$ tel que l'accroissement*

*$\Delta y$ en c vérifie l'inégalité*

$$\Delta y = f(x) - f(c) \leqslant 0, \quad \forall x \in ]c-\delta, c+\delta[$$
$$(resp.\ \Delta y = f(x) - f(c) \geqslant 0, \quad \forall x \in ]c-\delta, c+\delta[).$$

Le théorème de Fermat (voir § 12) nous dit que si une fonction $f$ a un extrémum local en un point $x_0$ et si la dérivée $f'(x_0)$ existe, alors

$$f'(x_0) = 0.$$

Par définition, un point $x_0$ est un *point stationnaire d'une fonction f* si $f'(x_0) = 0$.

Etant donnée une fonction $f$ sur un intervalle $]a, b[$, si l'on a à déterminer les points en lesquels elle présente un extrémum local, il faudra de toute évidence les chercher d'abord parmi les points stationnaires, c'est-à-dire parmi les points en lesquels la dérivée $f'$ est nulle et ensuite parmi les points (s'ils existent) en lesquels $f$ n'est pas dérivable. Les points stationnaires se déduisent à partir de l'équation

$$f'(x) = 0. \tag{1}$$

Signalons que la fonction $f$ n'a pas un extrémum local en tout point stationnaire.

La condition (1) est une condition *nécessaire* mais pas suffisante pour qu'une fonction dérivable $f$ ait un extrémum local en $x$. Par exemple, bien que $x = 0$ soit un point stationnaire pour la fonction $x^3$, celle-ci n'y présente pas d'extrémum local, car strictement croissante.

Il est évident aussi qu'une fonction $f$ n'a pas d'extrémum local en tout point où elle n'est pas dérivable.

De toute façon, si l'on sait qu'un point $x_0$ est stationnaire ou bien est un point en lequel la fonction $f$ n'admet pas de dérivée, il nous faut un critère nous permettant de dire si $f$ a bien un extrémum local en ce point et de déterminer la nature de cet extrémum.

On cite plus bas des conditions suffisantes d'extrémum local.

THÉORÈME 1. *Supposons que $x_0$ est un point stationnaire d'une fonction f (i.e. $f'(x_0) = 0$) et que f a une dérivée seconde continue dans un voisinage de $x_0$. Dans ces conditions*

*si $f''(x_0) < 0$, alors f a un maximum local en $x_0$;*
*si $f''(x_0) > 0$, alors f a un minimum local en $x_0$.*

DÉMONSTRATION. Développons la fonction $f$ en série de Taylor au voisinage de $x_0$ pour $n = 1$. Comme $f'(x_0) = 0$, la formule de Taylor de $f$ s'écrit

$$f(x) = f(x_0) + \frac{(x-x_0)^2}{2} f''(c), \qquad c \in ]x_0, x[. \tag{2}$$

Dans cette formule $x > x_0$ ou $x < x_0$.

Supposons que $f''(x_0) < 0$. La dérivée $f''$ étant continue au voisinage de $x_0$ par hypothèse, il existe un $\delta > 0$ tel que

$$f''(x) < 0, \quad \forall x \in ]x_0 - \delta, \ x_0 + \delta[.$$

Donc, dans la formule (2)

$$\frac{(x-x_0)^2}{2} f''(c) \leqslant 0, \quad \forall x \in ]x_0 - \delta, \ x_0 + \delta[,$$

ce qui montre que

$$\Delta y = f(x) - f(x_0) \leqslant 0, \quad \forall x \in ]x_0 - \delta, \ x_0 + \delta[,$$

c'est-à-dire que $f$ a un maximum local en $x_0$.

De façon analogue, si $f''(x_0) > 0$, alors $f''(x) > 0$ dans un voisinage de $x_0$ et $f''(c) > 0$. Donc, le reste de la formule (2) est positif au voisinage de $x_0$ et $\Delta y = f(x) - f(x_0) \geqslant 0$, c'est-à-dire que $f$ a un minimum local en $x_0$.

Exemple 1. $y = x^2 + 5$, $y' = 2x$, $x = 0$ est un point stationnaire ; $y'' = 2 > 0$ pour tous les $x$, donc en $x = 0$. Par conséquent, $y$ a un minimum local en $x = 0$.

Remarque 1. Si

$$f'(x_0) = 0 \text{ et } f''(x_0) = 0, \tag{3}$$

la fonction $f$ peut avoir ou non un extrémum en $x_0$. Ainsi, les fonctions $x^3$ et $x^4$ satisfont aux conditions (3) au point $x_0 = 0$, cependant la première n'a pas d'extrémum en 0, alors que la deuxième a un minimum.

Théorème 2. *Supposons que* $f'(x_0) = f''(x_0) = \ldots = f^{(n)}(x_0) = 0$ *et que* $f^{(n+1)}(x_0)$ *est différente de* 0 *en* $x_0$ *et continue dans un voisinage de* $x_0$. *Sous ces conditions*

*si* $n+1$ *est pair et* $f^{(n+1)}(x_0) < 0$, *alors* $f$ *a un maximum local en* $x_0$ ;

*si* $n+1$ *est pair et* $f^{(n+1)}(x_0) > 0$, *alors* $f$ *a un minimum local en* $x_0$ ;

*si* $n+1$ *est impair, alors* $f$ *n'a pas d'extrémum local en* $x_0$.

La démonstration de ce théorème repose de nouveau sur la formule de Taylor. On a

$$f(x) - f(x_0) = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(c), \quad c \in ]x_0, \ x[. \tag{4}$$

Si $n+1$ est pair, on raisonne comme pour le cas de la formule (2). Supposons maintenant que $n+1$ est impair. Par hypothèse, $f^{(n+1)}(x_0) \neq 0$. Etant continue au voisinage de $x_0$, la dérivée $f^{(n+1)}(x)$ conserve le signe de $f^{(n+1)}(x_0)$ dans un intervalle $]x_0 - \delta, \ x_0 + \delta[$. Si $x$ croît au voisinage de $x_0$, alors $(x - x_0)^{n+1}$ change de signe pour $x > x_0$ et $f^{(n+1)}(c)$ conservera son signe. Ceci montre que le second membre de (4) et partant $\Delta y = f(x) - f(x_0)$ change de signe pour $x > x_0$ et la fonction $f$ n'a pas d'extrémum en $x_0$.

THÉORÈME 3. *Supposons qu'une fonction $f$ est continue sur un intervalle $[x_0 - \delta, x_0 + \delta]$ et dérivable sur les intervalles $]x_0 - \delta, x_0[$ et $]x_0, x_0 + \delta[$. Ceci étant,*

$$f'(x) \geqslant 0 \ (\textit{resp.} \leqslant 0) \ \textit{sur} \ ]x_0 - \delta, x_0[, \tag{5}$$

$$f'(x) \leqslant 0 \ (\textit{resp.} \geqslant 0) \ \textit{sur} \ ]x_0, x_0 + \delta[. \tag{6}$$

*Alors la fonction $f$ a un maximum (resp. minimum) local en $x_0$.*

Ici l'existence de $f'(x_0)$ n'est pas indispensable.

DÉMONSTRATION. De la continuité de $f$ sur l'intervalle $[x_0 - \delta, x_0]$ et de la propriété (5) il s'ensuit (voir théorème 5 du § 12) que $f$

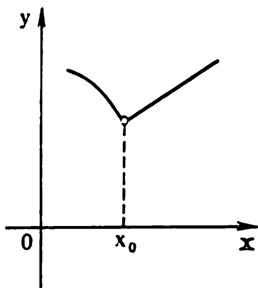



Fig. 52

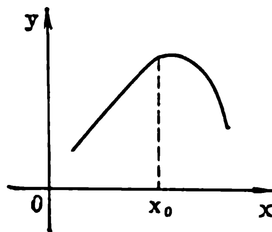



Fig. 53

est croissante (resp. décroissante) sur cet intervalle et par suite

$$f(x_0) - f(x) \geqslant 0 \ (\text{resp.} \leqslant 0) \ \text{pour} \ x \in [x_0 - \delta, x_0]. \tag{7}$$

De la continuité de $f$ sur $[x_0, x_0 + \delta]$ et de la propriété (6) il s'ensuit (voir théorème 5 du § 12) que

$$f(x) - f(x_0) \leqslant 0 \ (\text{resp.} \geqslant 0) \ \text{pour} \ x \in [x_0, x_0 + \delta]. \tag{8}$$

Il résulte alors de (7) et de (8) que:

$$f(x) \leqslant f(x_0) \ (\text{resp.} \ f(x) \geqslant f(x_0)) \ \forall x \in [x_0 - \delta, x_0 + \delta]$$

ce qui prouve le théorème 3.

Le théorème 3 affirme que si la dérivée première de la fonction $f$ change de signe en $x_0$, alors $f$ a un minimum en $x_0$ (fig. 52) si $f'$ passe du signe — au signe +, et un maximum (fig. 53) si elle passe du signe + au signe —. L'existence de $f'(x_0)$ n'est pas nécessaire, par contre la fonction $f$ doit être continue en $x_0$.

Etudier la fonction

$$f(x) = \begin{cases} 1 + x, & x < 0, \\ -x, & x \geqslant 0. \end{cases}$$

EXEMPLE 2. Soit la fonction $y = \frac{1}{1+x^2}$. Sa dérivée est $y' = \frac{-2x}{(1+x^2)^2}$.

On voit que $y' > 0$ pour $x < 0$, et $y' < 0$ pour $x > 0$, et de plus que $y$ est continue en $x = 0$. Donc, d'après le théorème 3 la fonction $y$ a un maximum local en $x = 0$. La fonction $y$ n'a pas d'autres extrémums locaux.

EXEMPLE 3. La fonction $y = 2 - x^2 (1 - \sin 1/x)$ $(x \neq 0)$, $y(0) = 2$, est continue en $x = 0$ et a un maximum local en $x = 0$: $y(x) \leqslant 2 = y(0)$. Cependant *on ne peut exhiber un voisinage du point* $x = 0$ dans lequel elle croît pour $x < 0$ et décroît pour $x > 0$. En effet,

$$y' = -2x\left(1 - \sin\frac{1}{x}\right) - \cos\frac{1}{x} \quad (x \neq 0),$$

$$y'(0) = \lim_{x \to 0} \frac{2 - x^2\left(1 - \sin\frac{1}{x}\right) - 2}{x} = -\lim_{x \to 0} x\left(1 - \sin\frac{1}{x}\right) = 0.$$

Pour $x$ petit, le terme $2x\left(1 - \sin\frac{1}{x}\right)$ est aussi petit que l'on veut, donc le signe de la dérivée $y'$ dépend de $\cos 1/x$. Lorsque $x \to 0$, $\cos 1/x$ prend les valeurs $\pm 1$ une infinité de fois. Donc, la fonction oscille dans tout voisinage du point $x = 0$.

THÉORÈME 4. *Si une fonction $f$ est telle que $f'(x_0) = 0$ et $f''(x_0) > 0$ (resp. $< 0$), alors elle a un minimum (resp. maximum) local en $x_0$.*

DÉMONSTRATION. Comme

$$f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f'(x)}{x - x_0} > 0,$$

alors $\frac{f'(x)}{x - x_0} > 0$ dans un voisinage assez petit de $x_0$, c'est-à-dire que $f'(x) < 0$ pour $x < x_0$ et $f'(x) > 0$ pour $x > x_0$. Le théorème 3 nous dit que $f$ a un minimum local en $x_0$. Le cas $f''(x_0) < 0$ se traite de la même manière.

REMARQUE 2. Le théorème 4 généralise le théorème 1, car il ne postule pas la continuité de $f''(x)$ au voisinage de $x_0$ mais seulement l'existence de $f''(x_0)$.

## § 18. Bornes d'une fonction sur un intervalle

Supposons qu'on ait à chercher le maximum (resp. minimum) d'une fonction $f$ continue sur un intervalle $[a, b]$. Que $f$ ait un maximum (resp. minimum) en un point $x_0 \in [a, b]$ a été prouvé dans le théorème 2 du § 5, chap. 3.

Trois cas seulement sont possibles: 1) $x_0 = a$, 2) $x_0 = b$, 3) $x_0 \in$ $\in ]a, b[$.

Si $x_0 \in ]a, b[$, alors d'après ce qui a été dit au § 17, la fonction $f$ a un extrémum local en $x_0$ qui est soit un point stationnaire, soit un point en lequel la dérivée n'existe pas.

Si de tels points forment un ensemble fini $\{x_1, \ldots, x_m\}$, alors

$$\max_{x\in[a,\,b]} f(x) = \max\{f(a),\ f(b),\ f(x_1),\ \ldots,\ f(x_m)\}$$

$$\text{(resp. } \min_{x\in[a,\,b]} f(x) = \min\{f(a),\ f(b),\ f(x_1),\ \ldots,\ f(x_m)\}\text{)}.$$

Signalons qu'il n'est pas nécessaire de connaître la nature des points stationnaires si l'on cherche seulement le maximum (resp. minimum) d'une fonction $f$ sur $[a,\ b]$.

EXEMPLE 1. Trouver le maximum et le minimum de la fonction

$$\psi(x) = \sin x + \cos x \text{ sur } [0,\ \pi].$$

Calculons la dérivée: $\psi'(x) = \cos x - \sin x$. Egalons-la à zéro:

$$\cos x - \sin x = 0.$$

Cette équation possède une seule racine $x = \pi/4$ sur l'intervalle $[0,\ \pi]$. Comme $\psi(0) = 1$, $\psi(\pi/4) = \sqrt{2}$, $\psi(\pi) = -1$, on a

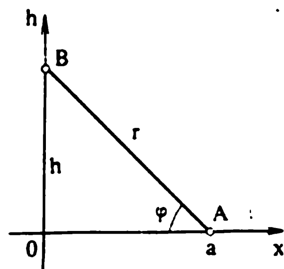



Fig. 54

$$\max_{x\in[0,\,\pi]} \psi(x) = \sqrt{2}, \quad \min_{x\in[0,\,\pi]} \psi(x) = -1.$$

EXEMPLE 2. A quelle hauteur $h$ faut-il suspendre une lampe électrique pour obtenir le meilleur éclairement en un point $A$ du plan non situé à la verticale de la lampe (fig. 54).

SOLUTION. Plaçons la lampe en $B$ et soient $AB = r$, $OB = h$, $OA = a$, $\widehat{OAB} = \varphi$. On sait que l'éclairement $I$ en $A$ est donné par la formule: $I = c\dfrac{\sin\varphi}{r^2}$, où $c$ est une constante. Traitons $h$ comme la variable. Comme $r^2 = h^2 + a^2$, $\sin\varphi = \dfrac{h}{r}$, il vient

$$I(h) = c\,\frac{h}{(h^2+a^2)^{3/2}}.$$

Dans la position du problème $h \in [0,\ \infty]$. Calculons le maximum de cette fonction. $I(0) = I(\infty) = 0$ [1]. D'autre part

$$I'(h) = c\,\frac{a^2-2h^2}{(h^2+a^2)^{5/2}} = 0 \text{ pour } h = a/\sqrt{2}.$$

Comme $I(a/\sqrt{2}) = 2c/3\sqrt{3}\,a^2 > 0$, la fonction $I(h)$ a son maximum au point $h = a/\sqrt{2}$.

Si la lampe est suspendue à une hauteur $h$ arbitraire, on aura le meilleur éclairement en un point $A$ tel que $a = \sqrt{2}h$.

[1]) Ici $I(\infty) = \lim\limits_{h\to+\infty} I(h)$.

## § 19. Convexité d'une courbe. Point d'inflexion

On dit qu'*une courbe $y = f(x)$ a sa convexité tournée vers le haut* (*resp. le bas*) *en $x_0$* si au voisinage de ce point elle est située au-dessous (resp. au-dessus) de la tangente en $x_0$ (sur la figure 55 la courbe a sa convexité tournée vers le bas en $x_1$ et vers le haut en $x_2$). Signalons qu'il y a équivalence entre convexité vers le haut (resp. le bas) et concavité vers le bas (resp. le haut).

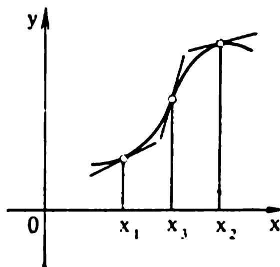


Fig. 55

On dit qu'un point $x_0$ est un *point d'inflexion de la courbe $y = f(x)$* si cette courbe traverse sa tangente en $x_0$ (sur la figure 55 le point $x_3$ est un point d'inflexion). Autrement dit, il existe un $\delta > 0$ assez petit pour que la courbe se trouve d'un côté de la tangente pour tous les $x \in ]x_0 - \delta, x_0[$ et de l'autre côté pour tous les $x \in ]x_0, x_0 + \delta[$.

Signalons que ces définitions n'épuisent pas toutes les dispositions de la courbe et de sa tangente en un voisinage assez petit du point de contact.

Par exemple, l'axe $Ox$ traverse et est tangent en $x = 0$ à la courbe

$$f(x) = \begin{cases} 0, & x = 0, \\ x^2 \sin \frac{1}{x}, & x \neq 0, \end{cases}$$

et $x = 0$ n'est pas un point d'inflexion.

THÉORÈME 1. *Si une fonction $f$ possède une dérivée seconde continue et $f''(x_0) > 0$ (resp. $< 0$), alors la courbe $y = f(x)$ a sa convexité tournée vers le bas (resp. le haut) en $x_0$.*

DÉMONSTRATION. Développons $f$ au voisinage de $x = x_0$ suivant la formule de Taylor

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r_1(x),$$

$$r_1(x) = \frac{(x - x_0)^2}{2} f''(x_0 + \theta(x - x_0)) \quad (0 < \theta < 1).$$

Formons l'équation de la tangente à la courbe $y = f(x)$ au point d'abscisse $x_0$:

$$Y = f(x_0) + f'(x_0)(x - x_0).$$

Le dépassement de la tangente en $x_0$ par la courbe $f$ est égal à

$$f(x) - Y = r_1(x).$$

La dérivée $f''$ étant continue, si $f''(x_0) > 0$, alors $f''(x_0 + \theta(x - x_0)) > > 0$ pour les $x$ appartenant à un voisinage assez petit de $x_0$, et par suite il est évident que $r_1(x) > 0$ pour tout $x \neq x_0$ de ce voisinage. Donc, le graphe de la fonction est situé au-dessus de la tangente et la courbe a sa convexité tournée vers le bas en $x_0$.

De façon analogue, si $f''(x_0) < 0$, alors $r_1(x) < 0$ pour tout $x \neq x_0$ d'un voisinage de $x_0$, c'est-à-dire que la courbe est située au-dessous de la tangente en $x_0$ et a sa convexité tournée vers le haut en $x_0$.

COROLLAIRE. *Si $x_0$ est un point d'inflexion d'une courbe $y = f(x)$ et si la dérivée seconde $f''$ existe en $x_0$, alors on a nécessairement $f''(x_0) = 0$.*

On cherchera donc les points d'inflexion d'une courbe $y = f(x)$ deux fois dérivable parmi les racines de l'équation $f''(x) = 0$.

La condition suffisante d'existence d'un point d'inflexion est donnée par le théorème suivant.

THÉORÈME 2. *Si une fonction $f$ est telle que sa dérivée $f'''$ est continue en $x_0$ et $f''(x_0) = 0$ et $f'''(x_0) \neq 0$, alors la courbe $y = f(x)$ a un point d'inflexion en $x_0$.*

DÉMONSTRATION. Dans ce cas

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r_2(x),$$
$$r_2(x) = \frac{(x - x_0)^3}{3!} f'''(x + \theta(x - x_0)).$$

De la continuité de $f'''$ en $x_0$ et du fait que $f'''(x_0) \neq 0$, il s'ensuit que $f'''(x_0 + \theta(x - x_0))$ conserve son signe dans un voisinage de $x_0$; ce signe est le même à gauche et à droite de $x_0$. D'autre part, le facteur $(x - x_0)^3$ change de signe en $x_0$ et avec lui $r_2(x)$. C.q.f.d.

Voici un théorème plus général.

THÉORÈME 3. *Soit une fonction $f$ telle que:*

$$f''(x_0) = \ldots = f^{(n)}(x_0) = 0,$$

*$f^{(n+1)}(x)$ est continue en $x_0$ et $f^{(n+1)}(x_0) \neq 0$.*

*Si $n$ est impair, alors la courbe $y = f(x)$ a sa convexité tournée vers le haut ou vers le bas selon que $f^{(n+1)}(x_0)$ est $< 0$ ou $> 0$.*

*Si $n$ est pair, alors $x_0$ est un point d'inflexion.*

La démonstration repose sur le fait que sous les conditions indiquées on a le développement suivant:

$$f(x) = f(x_0) + (x - x_0) f'(x_0) + \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(x_0 + \theta(x - x_0)).$$

Signalons en conclusion qu'on dit également qu'une courbe $y = f(x)$ a un point d'inflexion en un point $x$ où la dérivée $f'$ est infinie (fig. 40 et 41).

Par définition, une courbe $y = f(x)$ est *convexe vers le haut* (*resp. le bas*) *sur un intervalle* $[a, b]$ si tout arc de cette courbe d'extrémités $(x_1, f(x_1))$ et $(x_2, f(x_2))$ $(a \leqslant x_1 < x_2 \leqslant b)$ est situé au-dessus (resp. au-dessous) de la corde qui le sous-tend (fig. 56 et 57).

REMARQUE. Si $f$ est dérivable sur $[a, b]$, alors la définition de la convexité sur un intervalle est équivalente à la suivante: on dit qu'une courbe $y = f(x)$ est *convexe vers le haut* (*resp. le bas*) *sur un intervalle* $[a, b]$ si elle l'est en chaque point $x$ de l'intervalle $]a, b[$.

THÉORÈME 4. *Supposons qu'une fonction $f$ est continue sur $[a, b]$ et possède une dérivée seconde sur $]a, b[$.*

*Pour que la courbe $y = f(x)$ soit convexe vers le haut (resp. le bas) sur $[a, b]$, il est nécessaire et suffisant que $f''(x) \leqslant 0$ (resp. $f''(x) \geqslant 0$)*

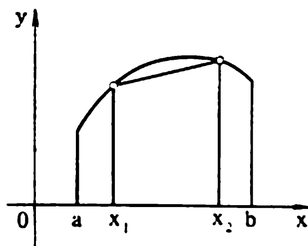



Fig. 56

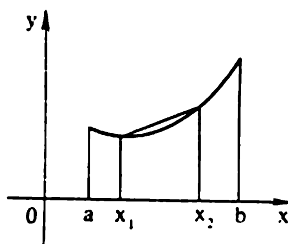



Fig. 57

*pour tous les $x \in ]a, b[$.*

Nous glisserons sur la démonstration de ce théorème.

EXEMPLE 1. La fonction $y = \sin x$ possède une dérivée première et une dérivée seconde continues: $(\sin x)'' = -\sin x \leqslant 0$ sur $[0,$

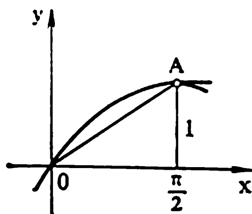



Fig. 58

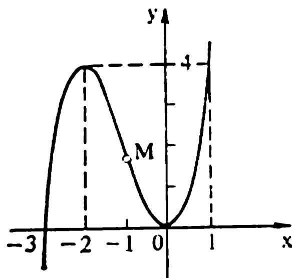



Fig. 59

$\pi/2]$. Donc, la corde $OA$ sous-tendant l'arc de courbe $y = \sin x$ sur $[0, \pi/2]$ est située au-dessous de la sinusoïde (fig. 58). L'équation de la corde étant $y = (2/\pi)x$, on obtient l'inégalité

$$\frac{2}{\pi} x \leqslant \sin x, \quad 0 \leqslant x \leqslant \pi/2,$$

qui est d'un usage courant en analyse.

EXEMPLE 2. $y = x^3 + 3x^2 = x^2(x + 3)$; $y' = 3x^2 + 6x$, $y' = 0$ pour $x = 0$, $x = -2$; $y'' = 6x + 6$, $y''(0) = 6 > 0$, $y''(-2) = -6 < 0$, $y'' = 0$ pour $x = -1$; $y''' = 6 \neq 0$. Comme $y'''(x) = 6 \neq 0$, le point $x = -1$ est un point d'inflexion. D'autre part, $y''(x) > 0$ pour $x > -1$, $y''(x) < 0$ pour $x < -1$. Donc, la courbe (fig. 59) est convexe vers le haut sur $]-\infty, -1[$ et vers le bas sur $]-1, \infty[$; en $x = 0$ on a un minimum et en $x = -2$, un maximum.

## § 20. Asymptote d'une courbe

On dit qu'une droite $x = a$ est une *asymptote verticale* d'une courbe continue $y = f(x)$ si l'une au moins des limites

$$\lim_{\substack{x \to a \\ x > a}} f(x), \quad \lim_{\substack{x \to a \\ x < a}} f(x)$$

est infinie.

Si une fonction $y = f(x)$ est définie pour $x > M$ (resp. $x < M$), on dit alors que la droite $Y = kx + b$ est une *asymptote oblique* de

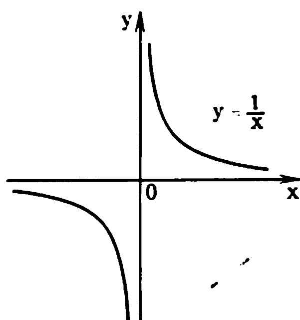



Fig. 60

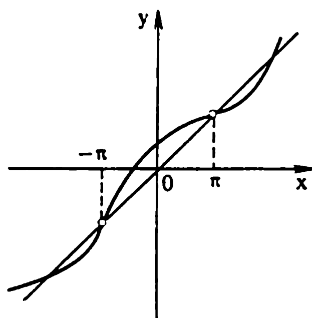



Fig. 61

la courbe continue $y = f(x)$ pour $x \to +\infty$ (resp. $x \to -\infty$) si

$$f(x) = kx + b + \alpha(x), \quad \text{où} \quad \lim_{\substack{x \to +\infty \\ (\text{resp. } x \to -\infty)}} \alpha(x) = 0$$

(autrement dit $|f(x) - kx - b|$ est un infiniment petit pour $x \to +\infty$ (resp. $x \to -\infty$)).

EXEMPLE 1. $y = 1/x$ (fig. 60); $x = 0$ est une asymptote verticale, car

$$\lim_{\substack{x \to 0 \\ x > 0}} \frac{1}{x} = +\infty, \quad \lim_{\substack{x \to 0 \\ x < 0}} \frac{1}{x} = -\infty.$$

EXEMPLE 2. $y = x + \frac{\sin x}{x}$ $(x \neq 0)$.

Comme $\lim\limits_{x \to \infty} \frac{\sin x}{x} = 0$, la droite $Y = x$ (fig. 61) est une asymptote oblique pour $x \to +\infty$ (et pour $x \to -\infty$).

EXEMPLE 3. $y = \sqrt{x}$ $(x \geqslant 0)$. Il est clair que $\sqrt{x} - kx - b$ ne tend pas vers 0 pour $x \to +\infty$ quels que soient $k$ et $b$, donc la courbe $y = \sqrt{x}$ ne possède pas d'asymptote oblique.

THÉORÈME. *Une condition nécessaire et suffisante pour qu'une courbe $y = f(x)$ admette une asymptote oblique pour $x \to +\infty$ $(x \to -\infty)$ est l'existence des limites finies*

$$\lim_{x \to +\infty} \frac{f(x)}{x} = k, \quad \lim_{x \to +\infty} [f(x) - kx] = b, \qquad (1)$$

*l'équation de cette asymptote étant $Y = kx + b$.*

DÉMONSTRATION. 1) Supposons que la courbe $y = f(x)$ admet une asymptote oblique d'équation $Y = kx + b$ pour $x \to +\infty$. Alors $f(x) = kx + b + \alpha(x)$, où $\alpha(x) \to 0$, $x \to +\infty$. D'où

$$\lim_{x \to +\infty} \frac{f(x)}{x} = \lim_{x \to +\infty} \left[k + \frac{b}{x} + \frac{\alpha(x)}{x}\right] = k,$$

$$\lim_{x \to +\infty} [f(x) - kx] = \lim_{x \to +\infty} [b + \alpha(x)] = b.$$

2) Supposons que les limites indiquées dans le théorème existent. De la deuxième relation on obtient par définition de la limite

$$f(x) - kx - b = \alpha(x), \text{ où } \alpha(x) \to 0 \text{ pour } x \to +\infty,$$

c'est-à-dire que $f(x) = kx + b + \alpha(x)$. Donc, la droite $Y = kx + b$ est une asymptote oblique pour $x \to +\infty$. On raisonne de façon analogue pour $x \to -\infty$.

Si $k = 0$, l'*asymptote* est dite *horizontale*.

REMARQUE. L'existence des limites *finies* (1) est essentielle. En effet, pour la courbe $y = \sqrt{x}$ $(x \geqslant 0)$ on a $\lim\limits_{x \to +\infty} \frac{\sqrt{x}}{x} = 0 = k$ et

$$\lim_{x \to +\infty} [\sqrt{x} - 0 \cdot x] = \infty, \text{ c'est-à-dire que } b = \infty$$

et cette courbe n'a pas d'asymptote.

Voici une définition équivalente de l'asymptote oblique.

*On dit qu'une droite $y = kx + b$ est une asymptote oblique d'une courbe $y = f(x)$ pour $x \to \pm\infty$ si la distance $\rho(x)$ d'un point $A(x, f(x))$ de la courbe à la droite tend vers 0 pour $x \to \pm\infty$.*

En effet, on sait que la distance d'un point $(x, f(x))$ à la droite $y = kx + b$ s'exprime par la formule

$$\rho(x) = |f(x) - kx - b| / \sqrt{1 + k^2},$$

d'où il s'ensuit, puisque $|f(x) - kx - b| \to 0$, que $\rho(x) \to 0$, et réciproquement.

Exemple 4. Etablir l'existence des asymptotes de l'hyperbole

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (|x| \geqslant a,\ a \geqslant b > 0).$$

En résolvant cette équation par rapport à $y$, on obtient

$$y = \pm \frac{b}{a} \sqrt{x^2 - a^2}.$$

D'où

$$\lim_{x \to +\infty} \frac{y}{x} = \pm \frac{b}{a} \lim_{x \to +\infty} \frac{\sqrt{x^2 - a^2}}{x} = \pm \frac{b}{a}.$$

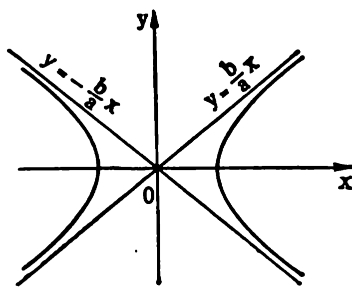



Fig. 62



D'autre part

$$\lim_{x \to +\infty} \left[ y - \left( \pm \frac{b}{a} x \right) \right] =$$

$$= \pm \frac{b}{a} \lim_{x \to +\infty} [\sqrt{x^2 - a^2} - x] = \pm \frac{b}{a} \lim_{x \to +\infty} \frac{-a^2}{\sqrt{x^2 - a^2} + x} = 0.$$

Donc, le théorème prouvé nous dit que les droites

$$y = \pm \frac{b}{a} x$$

sont les asymptotes de l'hyperbole, le signe + correspondant à la demi-branche supérieure de droite, le signe — à la demi-branche inférieure de droite.

Pour des raisons de symétrie, il est évident que ces droites sont asymptotes pour $x \to -\infty$. Dans ce cas le signe + correspond à la demi-branche inférieure de gauche, le signe —, à la demi-branche supérieure de gauche (fig. 62).

## § 21. Courbe continue et lisse

Les équations

$$\begin{cases} x = \varphi(t), \\ y = \psi(t) \end{cases} \quad (a < t < b), \tag{1}$$

où $\varphi$ et $\psi$ sont des fonctions continues sur $]a, b[$, définissent une *courbe continue à l'aide d'un paramètre t* ou encore le lieu géométrique des points de coordonnées $(\varphi(t), \psi(t))$, $t \in ]a, b[$.

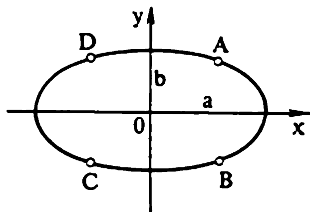


Fig. 63

La courbe continue (1) est par définition *lisse* si $\varphi(t)$ et $\psi(t)$ ont des dérivées continues sur $]a, b[$ et

$$(\varphi'(t))^2 + (\psi'(t))^2 > 0, \quad \forall t \in ]a, b[. \tag{2}$$

Désignons la courbe (1) par $\Gamma$. Soit $t_0 \in ]a, b[$. D'après (2), l'un des nombres $\varphi'(t_0)$ et $\psi'(t_0)$ est différent de 0. Supposons pour fixer les idées que $\varphi'(t_0) \neq 0$. La dérivée $\varphi'(t)$ étant continue, il existe un intervalle $]t_0 - \delta, t_0 + \delta[$ sur lequel $\varphi'(t)$ est du signe de $\varphi'(t_0)$. Donc, $\varphi(t)$ est strictement monotone sur $[t_0 - \delta, t_0 + \delta]$. On sait d'autre part qu'elle est continûment dérivable. Donc, la fonction $x = \varphi(t)$ admet une réciproque

$$t = \varphi^{-1}(x) = g(x), \tag{3}$$

strictement monotone et continûment dérivable sur un intervalle $]c, d[$ qui est un voisinage du point $x_0 = \varphi(t_0)$.

En portant l'expression de $t$ dans la deuxième équation (1), on obtient une fonction continûment dérivable (voir théorème du § 4)

$$y = F(x) = \psi[\varphi^{-1}(x)], \quad x \in ]c, d[, \tag{4}$$

représentée par une portion $\gamma$ de la courbe $\Gamma$ correspondant à l'intervalle $]t_0 - \delta, t_0 + \delta[$. Donc, $\gamma$ admet en tout point une tangente non parallèle à l'axe $Oy$. Il est évident que les points de $\gamma$ se projettent de façon unique sur l'axe $Ox$.

Si maintenant $\psi'(t_0) \neq 0$, on obtient par des raisonnements analogues une fonction continûment dérivable

$$x = \Phi(y) = \varphi[\psi^{-1}(y)], \quad y \in ]c_1, d_1[, \tag{5}$$

représentée graphiquement par une portion $\gamma_1$ de la courbe $\Gamma$ correspondant à l'intervalle $]t_0 - \delta, t_0 + \delta[$. D'où il résulte que $\gamma_1$ admet en tout point une tangente qui n'est pas parallèle à l'axe $Ox$.

Donc, la courbe lisse Γ possède en chacun de ses points une tangente qui est susceptible d'être parallèle à l'un des axes de coordonnées.

Exemple. Les équations

$$\begin{cases} x = a\cos t, \\ y = b\sin t, \end{cases} \quad t \in ]-\infty, \infty[,$$

sont les équations paramétriques d'une ellipse d'axes $2a$ et $2b$ (fig. 63).

C'est une courbe lisse, car les fonctions $x = a\cos t$ et $y = b\sin t$ possèdent des dérivées continues non simultanément nulles:

$$(x'(t))^2 + (y'(t))^2 = (a\sin t)^2 + (b\cos t)^2 \geqslant$$
$$\geqslant b^2(\sin^2 t + \cos^2 t) = b^2 > 0 \ (0 < b \leqslant a).$$

Les points $A$, $B$, $C$ et $D$ partagent l'ellipse en quatre portions lisses se projetant chacune de façon unique sur l'axe $Ox$ ou sur l'axe $Oy$.

## § 22. Construction du graphe d'une fonction

Pour construire le graphe d'une fonction $y = f(x)$, on doit procéder de la manière suivante.

1. Déterminer l'ensemble de définition et étudier la continuité.
2. Chercher les symétries et les périodes afin de limiter l'ensemble où l'on étudie la fonction.
3. Déterminer le sens de variation. Pour cela chercher les points où la dérivée $f'(x)$ est nulle ou n'existe pas ou est égale à $\infty$. Calculer les valeurs de $f$ en ces points (si elles existent), puis s'assurer de l'existence d'un extrémum en points.
4. Etudier la fonction aux bornes des intervalles où elle est définie et continue: si l'intervalle de définition est fermé en une extrémité $a$, on calcule $f(a)$; si cet intervalle est ouvert en une extrémité $b$, où $+\infty$ ou $-\infty$, on cherche si $f(x)$ possède une limite quand $x$ tend vers cette extrémité.

Les renseignements obtenus permettent de dresser le tableau de variation. Pour obtenir la courbe représentative:

5. On détermine les asymptotes en calculant les limites

$$\lim_{x\to\pm\infty} \frac{f(x)}{x} = k \quad \text{et} \quad \lim_{x\to\pm\infty} [f(x) - kx] = b$$

si elles existent.

6. On cherche les points remarquables de la courbe et éventuellement leurs tangentes: points où la tangente est parallèle à l'axe $Ox$ (ces points sont obtenus en cherchant les valeurs qui annulent la dérivée), points d'arrêt, points anguleux, points d'intersection avec les asymptotes ou les axes de coordonnées. On cherchera aussi les points en lesquels $f''(x) = 0$. Ces points peuvent être des points d'inflexion. On étudie le signe de $f''(x)$ pour déterminer la convexité de la courbe.

Consigner enfin toutes ces données dans un tableau de la forme suivante :

| $x$ | $-\infty$ | | 0 | | $\frac{1}{2}$ | | 1 | | $\infty$ |
|---|---|---|---|---|---|---|---|---|---|
| $y'$ | | + | 0 | − | $-\frac{3}{2}$ | − | 0 | + | |
| $y$ | | ↗ | 0 | ↘ | $-\frac{1}{2}$ | ↘ | −1 | ↗ | |
| | pas d'asymptote | | | | | | | asymptote $y=x-3$ | |
| $y''$ | | − | | − | 0 | + | | + | |
| $y$ | | convexité vers le haut | max | convexité vers le haut | inflexion | convexité vers le bas | min | convexité vers le bas | |

La courbe de la fonction étudiée dans le tableau précédent est représentée sur la figure 64.

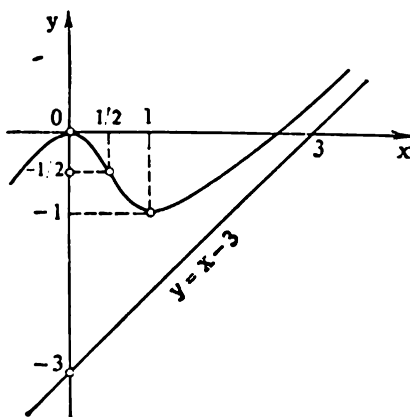



Fig. 64

Ce graphe ne nous donne certes les valeurs exactes de la fonction qu'aux points $x = 0$, $1/2$, $1$, les autres valeurs étant prises au jugé ; néanmoins son mérite est de nous donner une idée sur l'allure géné-

rale de la fonction. Si on veut la tabuler sur un certain intervalle pour des valeurs de $x$ distantes de 0,001 par exemple, il faudrait alors recourir à d'autres instruments de calcul (calculette, calculatrice ou autre). Dans tous les cas, il est profitable de connaître l'allure générale de la courbe.

EXEMPLE. Tracer la courbe définie paramétriquement par les équations

$$\begin{cases} x = te^t, \\ y = te^{-t}, \end{cases} \quad t \in ]-\infty, \infty[. \tag{1}$$

SOLUTION. Traçons d'abord la courbe représentative de la fonction $x = te^t$.

Ensemble de définition: $x$ est définie et continue dans l'intervalle $]-\infty, \infty[$.

Sens de variation: $x' = e^t + te^t = (1 + t)\, e^t$. L'équation $x'(t) = 0$ possède une racine unique $t = -1$. Il est évident que $x' > 0$ pour $t > -1$ et $x' < 0$ pour $t < -1$. Donc, la fonction $x(t)$ est strictement croissante pour $t > -1$ et strictement décroissante pour $t < -1$.

Etude aux bornes: lorsque $t \to -\infty$, $x \to 0$,
lorsque $t \to \infty$, $x \to \infty$.

**Tableau de variation**

| $t$ | $-\infty$ | | $-1$ | | $\infty$ |
|---|---|---|---|---|---|
| $x'$ | | $-$ | $0$ | $+$ | |
| $x$ | $0$ | $\searrow$ | $-e^{-1}$ | $\nearrow$ | $\infty$ |

Asymptotes:

$$\lim_{t \to -\infty} \frac{te^t}{t} = 0, \quad \lim_{t \to -\infty} [te^t - 0] = 0,$$

donc $x = 0$ est une asymptote horizontale.

Points remarquables. Le tableau de variation met en évidence l'existence d'un minimum local au point $(-1, -e^{-1})$.

Point à tangente parallèle à $Ot$: $(-1, -e^{-1})$.

Point d'inflexion: $x'' = (2 + t)\, e^t$. Donc,

$$x'' > 0 \text{ pour } t > -2,$$
$$x'' < 0 \text{ pour } t < -2,$$
$$x''(-2) = 0.$$

La courbe $x(t)$ présente une inflexion en $t = -2$, est convexe vers

le haut sur $]-\infty, -2[$, convexe vers le bas sur $]-2, \infty[$. Le graphe est représenté sur la figure 65.

Construisons maintenant le graphe de la fonction $y = te^{-t}$.

Ensemble de définition : $y$ est définie et continue sur $]-\infty, \infty[$.

Sens de variation : $y' = (1 - t)\, e^{-t}$. L'équation $y'(t) = 0$ possède une racine unique $t = +1$. Il est évident que $y' > 0$ pour $t < 1$ et $y' < 0$ pour $t > 1$. Donc, la fonction $y(t)$ est strictement croissante pour $t < 1$ et strictement décroissante pour $t < 1$.

Etude aux bornes : lorsque $t \to \infty$, $y(t) \to 0$,
lorsque $t \to -\infty$, $y(t) \to -\infty$.

**Tableau de variation**

| $t$ | $-\infty$ | | $1$ | | $\infty$ |
|---|---|---|---|---|---|
| $y'$ | | $+$ | $0$ | $-$ | |
| $y$ | $-\infty$ | $\nearrow$ | $e^{-1}$ | $\searrow$ | $0$ |

Asymptotes :

$$\lim_{t\to\infty} \frac{te^{-t}}{t} = 0, \quad \lim_{t\to\infty} [te^{-t} - 0] = 0,$$

donc $y = 0$ est une asymptote horizontale.

Points remarquables. Le tableau met en évidence l'existence d'un maximum local en $t = 1$.

Point à tangente parallèle à $Ot$ : $(1.\ e^{-1})$.

Point d'inflexion : $y'' = (t - 2)\, e^{-t}$. Donc,

$$y'' > 0 \text{ pour } t > 2,$$
$$y'' < 0 \text{ pour } t < 2,$$
$$y'' = 0 \text{ pour } t = 2.$$

Donc, $y(t)$ présente une inflexion en $t = 2$, est convexe vers le haut pour $t \in ]-\infty, 2[$ et vers le bas pour $t \in ]2, \infty[$. Le graphe est repré-

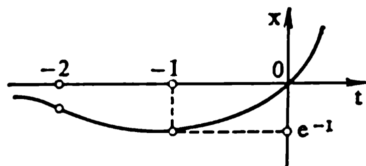


Fig. 65

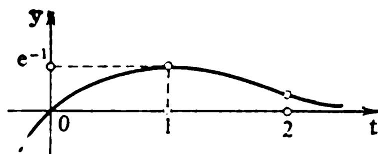


Fig. 66

senté sur la figure 66.

Passons maintenant à un problème plus compliqué, au tracé de

la courbe (1). Désignons-la par $\Gamma$. Les fonctions qui la définissent sont continûment dérivables autant de fois qu'on le veut. Nous utiliserons seulement le fait qu'elles sont bicontinûment dérivables. Signalons que $\Gamma$ est une courbe lisse, car les dérivées de $x = \varphi(t) = te^t$ et de $y = \psi(t) = te^{-t}$ ne sont pas simultanément nulles.

Soient $\Gamma_1$ et $\Gamma_2$ les branches respectives de $\Gamma$ sur lesquelles $x'_t < 0$ et $x'_t > 0$. Donc (voir fig. 65 et 66),

$\Gamma_1$ correspond aux variations de $t$ sur $]-\infty, -1[$,

$\Gamma_2$, aux variations de $t$ sur $]-1, \infty[$.

La fonction $x = \varphi(t)$ représentée graphiquement par $\Gamma_1$ est strictement décroissante de $\varphi(-\infty) = 0$ à $\varphi(-1) = -e^{-1}$. Elle admet

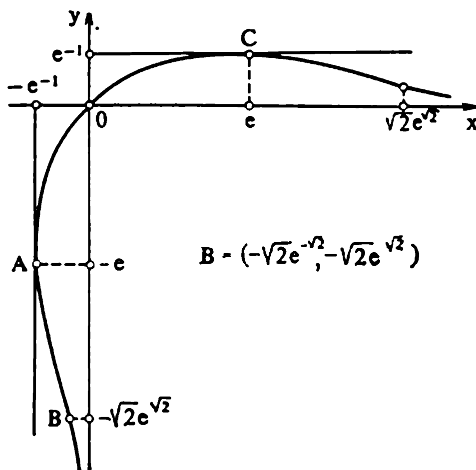



Fig. 67

donc une réciproque. La fonction $y = \psi(t)$ croît strictement de $\psi(-\infty) = -\infty$ à $\psi(-1) = -e$. Il s'ensuit que $\Gamma_1$ est la courbe représentative de la fonction explicite

$$y = \psi[\varphi^{-1}(x)], \quad x \in ]-e^{-1}, 0[.$$

$\Gamma_1$ est représentée sur la figure 67 en-dessous du point $A$. Lorsque $t$ passe de $-\infty$ à $-1$, l'abscisse $x$ du point courant de $\Gamma_1$ passe de 0 à $-e^{-1}$, et l'ordonnée $y$, de $-\infty$ à $-e$. Comme $x'(-1) = 0$ et $y'(1) \neq \neq 0$, la tangente en $A$ est parallèle à $Oy$. Les points de $\Gamma$ sont tous situés à droite de la tangente, car $x \geqslant -e^{-1}$ (voir fig. 65).

En tout point de $\Gamma$ autre que $A$, c'est-à-dire tel que $t \neq -1$, la dérivée $x'(t) \neq 0$ et

$$y'_x = \frac{y'_t}{x'_t} = \frac{1-t}{1+t} e^{-2t}, \tag{2}$$

$$y''_x = \frac{dy'_x}{dx} = \frac{\left(\frac{1-t}{1+t}e^{-2t}\right)'}{(1+t)e^t} = 2\frac{t^2-2}{(1+t)^3}e^{-3t}. \qquad (3)$$

D'où

$$y''_x\,|_{t=\pm\sqrt{2}} = 0. \qquad (4)$$

Etudions le signe de $y''_x$:

| $t$ | $-\infty$ | | $-\sqrt{2}$ | | $-1$ | | $\sqrt{2}$ | | $\infty$ |
|---|---|---|---|---|---|---|---|---|---|
| $y''_x$ | | $-$ | 0 | $+$ | ‖ | $-$ | 0 | $+$ | |

Donc, $\Gamma_1$ a une inflexion au point $B(-\sqrt{2}e^{-\sqrt{2}}, -\sqrt{2}e^{\sqrt{2}})$, est convexe vers le haut pour $t < -\sqrt{2}$ et vers le bas pour $t \in$ $\in ]-\sqrt{2}, -1[$.

Passons maintenant à $\Gamma_2$ $(-1 < t < \infty)$. On voit sur les figures 65 et 66 que les fonctions $x = \varphi(t)$ et $y = \psi(t)$ sont strictement croissantes sur l'intervalle $]-1, 1[$, donc il en est de même de la fonction

$$y = \psi[\varphi^{-1}(x)]$$

sur l'intervalle $]-e^{-1}, e[$. Sur cet intervalle la courbe a sa convexité tournée vers le haut (en vertu de (3)). Au point $C$, on a $y'_x = 0$ $\left(y'_x(e) = \frac{y'(1)}{x'(1)} = \frac{0}{x'(1)} = 0\right)$, donc la fonction $y(x)$ présente un maximum local en $C$, car convexe vers le haut. Lorsque $x > e$ (i.e. $t > 1$), $x(t)$ tend vers $\infty$ et $y(t)$, vers 0, c'est-à-dire que $y(x) \underset{x\to\infty}{\to} 0$ par valeurs décroissantes. On voit sur (3) que le deuxième point d'inflexion de $y(x)$ est le point de coordonnées $(\sqrt{2}e^{\sqrt{2}}, \sqrt{2}e^{-\sqrt{2}})$. A gauche de ce point, la courbe est convexe vers le haut, à droite, vers le bas.

## § 23. Fonction vectorielle. Vecteurs tangent et normal

Soit un plan rapporté à un système de coordonnées rectangulaires $(Ox, Oy)$. Les équations

$$\begin{cases} x = x(t), \\ y = y(t), \end{cases} \quad t \in ]a, b[, \qquad (1)$$

où $x(t)$ et $y(t)$ sont des fonctions continues sur $]a, b[$, définissent une *courbe continue* $\Gamma$, lieu géométrique des points dont les coordonnées $x$ et $y$ vérifient (1). L'équation de $\Gamma$ peut être représentée sous la

forme vectorielle

$$\boldsymbol{r}(t) = x(t)\,\boldsymbol{i} + y(t)\,\boldsymbol{j}, \quad t \in ]a, b[, \tag{1'}$$

où $\boldsymbol{i}$ et $\boldsymbol{j}$ sont les vecteurs unités de $Ox$ et de $Oy$ respectivement, $\boldsymbol{r} = \boldsymbol{r}(t)$ le rayon vecteur du point de $\Gamma$ correspondant à la valeur $t$ du paramètre (fig. 68).

Le vecteur $\boldsymbol{r}(t)$ s'appelle *fonction vectorielle* (définie sur $]a, b[$). On dit de ce fait que la courbe $\Gamma$ est l'*hodographe* de la fonction

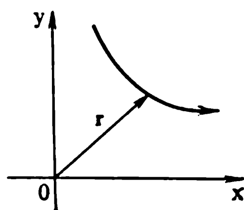



Fig. 68

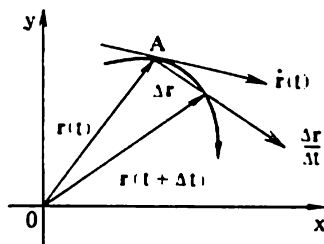



Fig. 69

vectorielle $\boldsymbol{r}(t)$, c'est-à-dire le lieu géométrique des extrémités des vecteurs $\boldsymbol{r}(t)$ issus de $O$.

On dit que la courbe $\Gamma$ est *lisse* sur $]a, b[$ si les fonctions $x(t)$ et $y(t)$ ont des dérivées continues sur $]a, b[$ non simultanément nulles.

A un accroissement $\Delta t$ de $t$ correspond l'accroissement $\Delta\boldsymbol{r}$ (fig. 69) :

$$\Delta\boldsymbol{r} = \boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t) = [x(t+\Delta t) - x(t)]\,\boldsymbol{i} + \\ + [y(t+\Delta t) - y(t)]\,\boldsymbol{j} = \Delta x\boldsymbol{i} + \Delta y\boldsymbol{j},$$

d'où, en divisant scalairement par $\Delta t$, on obtient

$$\frac{\Delta\boldsymbol{r}}{\Delta t} = \frac{\Delta x}{\Delta t}\boldsymbol{i} + \frac{\Delta y}{\Delta t}\boldsymbol{j}.$$

Pour toute courbe lisse on a

$$\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t} = x', \quad \lim_{\Delta t\to 0}\frac{\Delta y}{\Delta t} = y'.$$

Le vecteur $x'\boldsymbol{i} + y'\boldsymbol{j}$ s'appelle *dérivée de* $\boldsymbol{r}$ en $t$ et se note

$$\dot{\boldsymbol{r}} = x'\boldsymbol{i} + y'\boldsymbol{j}.$$

On peut définir encore la dérivée $\dot{\boldsymbol{r}}$ comme un vecteur tel que

$$\left|\frac{\Delta\boldsymbol{r}}{\Delta t} - \dot{\boldsymbol{r}}\right| \to 0, \quad \Delta t \to 0.$$

En effet

$$\left|\frac{\Delta\boldsymbol{r}}{\Delta t} - \dot{\boldsymbol{r}}\right|^2 = \left(\frac{\Delta x}{\Delta t} - x'\right)^2 + \left(\frac{\Delta y}{\Delta t} - y'\right)^2 \to 0, \quad \Delta t \to 0.$$

On dit que le vecteur $\dot{r}$ est la *limite du vecteur* $\Delta r/\Delta t$ *pour* $\Delta t \to 0$ et on note

$$\dot{r} = \lim_{\Delta t \to 0} \frac{\Delta r}{\Delta t}.$$

Sur la figure 69, on voit que le vecteur $\dot{r}$ est dirigé suivant la tangente à $\Gamma$ en $t$ et orienté dans le sens des $t$ croissants.

Le vecteur $\dot{r}$ s'appelle *vecteur tangent* à $\Gamma$. Son module est

$$|\dot{r}| = \sqrt{x'^2 + y'^2}.$$

Le *vecteur unité tangent* est

$$\tau = \frac{\dot{r}}{|\dot{r}|} = \cos\alpha i + \sin\alpha j \qquad (|\dot{r}| > 0),$$

$$\cos\alpha = \frac{x'}{\sqrt{x'^2+y'^2}}, \quad \sin\alpha = \frac{y'}{\sqrt{x'^2+y'^2}}, \tag{2}$$

où $\alpha$ est l'angle que fait $\tau$ avec $Ox$.

Le *vecteur unité normal* à $\Gamma$, c'est-à-dire le vecteur unité orthogonal à $\tau$, est défini par

$$\nu = (\nu_1, \nu_2),$$

$$\nu_1 = \mp \sin\alpha, \ \nu_2 = \pm \cos\alpha \tag{3}$$

ou

$$\nu_1 = \mp \frac{y'}{\sqrt{x'^2+y'^2}}, \quad \nu_2 = \pm \frac{x'}{\sqrt{x'^2+y'^2}}. \tag{3'}$$

Le déterminant

$$\begin{vmatrix} \tau_1 & \tau_2 \\ \nu_1 & \nu_2 \end{vmatrix} = \begin{vmatrix} \cos\alpha & \sin\alpha \\ \mp\sin\alpha & \pm\cos\alpha \end{vmatrix} = \pm 1.$$

Les signes supérieurs correspondent au cas où le couple de vecteurs $(\tau, \nu)$ est orienté comme le couple $(i, j)$ (fig. 70), les signes inférieurs, au cas où le couple $(\tau, \nu)$ est orienté dans le sens contraire (fig. 71).

La dérivée seconde de la fonction vectorielle $r(t)$ (voir (1')) se définit comme la limite

$$\ddot{r}(t) = \lim_{\Delta t \to 0} \frac{\dot{r}(t+\Delta t) - \dot{r}(t)}{\Delta t} = x''(t)\, i + y''(t)\, j.$$

La figure 72 représente une courbe $\Gamma$; le point $A$ correspond à une valeur $t$, le point $B$, à la valeur $t + \Delta t$. En ces points on a tracé les vecteurs tangents $\dot{r}(t)$ et $\dot{r}(t + \Delta t)$, puis on a transporté le vecteur $\dot{r}(t + \Delta t)$ parallèlement en $A$. On a construit le vecteur

$\Delta\dot{r} = \dot{r}(t + \Delta t) - \dot{r}(t)$ et le vecteur $\Delta\dot{r}/\Delta t$ de même sens que $\Delta\dot{r}$. On a enfin porté le vecteur limite $\ddot{r} = \ddot{r}(t)$. *Le vecteur* $\ddot{r}$ *est orienté dans le sens de la concavité de* $\Gamma$. De façon plus précise : *le vecteur* $\ddot{r}$

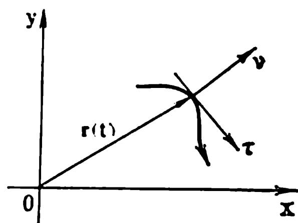



Fig. 70

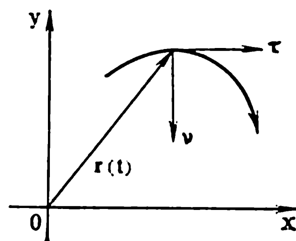



Fig. 71

*fait un angle aigu avec le vecteur normal* $\boldsymbol{\nu}$ *à* $\Gamma$ *qui est orienté dans le sens de la concavité de* $\Gamma$.

EXEMPLE. L'équation vectorielle de l'ellipse (voir § 21) est

$$r = ia \cos t + jb \sin t \quad (t \in ]-\infty, \infty[).$$

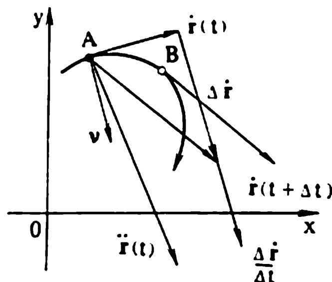



Fig. 72

Le vecteur tangent est

$$\dot{r} = -ia \sin t + jb \cos t,$$

le vecteur normal

$$n = \mp ib \cos t \mp ja \sin t.$$

Ici $n$ n'est pas en général un vecteur unité.

La fonction vectorielle $r = r(t)$ peut être développée au voisinage d'un point $t_0$ en une série vectorielle de Taylor. Soit

$$r(t) = x(t)\, i + y(t)\, j,$$

où $x(t)$ et $y(t)$ sont dérivables autant de fois qu'on le veut au voisinage de $t_0$. En développant ces fonctions d'après la formule de Taylor, on obtient

$$x(t)=x(t_0)+\frac{x'(t_0)}{1!}(t-t_0)+\ldots+\frac{x^{(n)}(t_0)}{n!}(t-t_0)^n+R_n(t), \quad (4)$$

$$y(t)=y(t_0)+\frac{y'(t_0)}{1!}(t-t_0)+\ldots+\frac{y^{(n)}(t_0)}{n!}(t-t_0)^n+\overline{R}_n(t), \quad (5)$$

où $R_n(t)$ et $\overline{R}_n(t)$ sont les restes (de Lagrange ou de Cauchy ou dans une autre forme). En multipliant (4) par $\boldsymbol{i}$ et (5) par $\boldsymbol{j}$ et en ajoutant, on obtient la formule de Taylor de la fonction vectorielle $\boldsymbol{r}(t)$:

$$\boldsymbol{r}(t)=\boldsymbol{r}(t_0)+\frac{\boldsymbol{r}'(t_0)}{1!}(t-t_0)+\ldots+\frac{\boldsymbol{r}^{(n)}(t_0)}{n!}(t-t_0)^n+\boldsymbol{r}_n(t),$$

où le reste

$$\boldsymbol{r}_n(t)=R_n(t)\,\boldsymbol{i}+\overline{R}_n(t)\,\boldsymbol{j}.$$

Signalons que si les restes $R_n(t)$ et $\overline{R}_n(t)$ sont écrits sous la forme de Lagrange ou de Cauchy, les dérivées $x^{(n+1)}(t)$ et $y^{(n+1)}(t)$ se calculent généralement en des points différents.

CHAPITRE 5

# INTÉGRALES INDÉFINIES

## § 1. Intégrale indéfinie. Table des intégrales

Dans le chapitre précédent nous avons présenté la notion de dérivée et appris à calculer les dérivées des fonctions élémentaires. Nous nous proposons maintenant de résoudre le problème inverse, c'est-à-dire trouver une fonction $f(x)$ dont on connaît la dérivée $f'(x)$.

En mécanique par exemple cela revient à déterminer la loi du mouvement d'un point matériel dont on connaît la vitesse.

DÉFINITION. *On dit qu'une fonction $F(x)$ est une primitive d'une fonction $f(x)$ sur un intervalle $]a, b[$ si $F(x)$ est dérivable sur $]a, b[$ et $F'(x) = f(x)$.*

On définit de façon analogue la notion de primitive sur un intervalle $[a, b]$ en considérant les dérivées unilatérales en $a$ et en $b$.

EXEMPLE 1. $F(x) = \sqrt{x}$ est une primitive de la fonction $f(x) = \frac{1}{2\sqrt{x}}$ sur $]0, \infty[$, puisque $(\sqrt{x})' = \frac{1}{2\sqrt{x}}$.

EXEMPLE 2. $F(x) = \sin 2x$ est une primitive de la fonction $f(x) = 2\cos 2x$ sur $]-\infty, \infty[$, car $(\sin 2x)' = 2\cos 2x$.

THÉORÈME 1. *Si $F(x)$ est une primitive de la fonction $f(x)$ sur $]a, b[$, alors $F(x) + C$, où $C$ est une constante, l'est également.*

DÉMONSTRATION. On a $(F(x) + C)' = F'(x) = f(x)$.

THÉORÈME 2. *Si $F_1(x)$ et $F_2(x)$ sont des primitives de $f(x)$ sur $]a, b[$, alors $F_1(x) - F_2(x) = C$, où $C$ est une constante.*

DÉMONSTRATION. Par hypothèse $F_1'(x) = F_2'(x) = f(x)$. Formons la fonction $\Phi(x) = F_1(x) - F_2(x)$. Il est évident que

$$\Phi'(x) = F_1'(x) - F_2'(x) = f(x) - f(x) = 0, \quad \forall\, x \in ]a, b[.$$

De là, on déduit en vertu du théorème 6 (chap. 4, § 12) que $\Phi(x) \equiv C$, autrement dit $F_1(x) - F_2(x) = C$, c.q.f.d.

Donc, des théorèmes 1 et 2 il résulte que si $F(x)$ est une primitive de $f(x)$ sur $]a, b[$, toute autre primitive $\Phi(x)$ de $f(x)$ sur $]a, b[$ est de la forme

$$\Phi(x) = F(x) + C, \tag{1}$$

où $C$ est une constante (fig. 73).

Définition. *On appelle intégrale indéfinie d'une fonction $f(x)$ et on note*

$$\int f(x)\,dx \tag{2}$$

*toute primitive de $f(x)$ sur $]a, b[$. Le signe $\int$ s'appelle intégrale, $f(x)\,dx$ est l'intégrant, $f(x)$, la fonction à intégrer.*

Si $F(x)$ est une primitive de $f(x)$, alors d'après ce qui précède

$$\int f(x)\,dx = F(x) + C, \tag{3}$$

où $C$ est une constante dûment choisie.

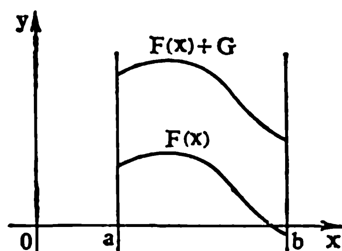



Fig. 73

L'opération qui consiste à trouver l'intégrale indéfinie d'une fonction $f(x)$ s'appelle *intégration* de $f(x)$.

Signalons que si $F(x)$ est une primitive de $f(x)$, alors l'intégrant $f(x)dx = F'(x)\,dx = dF(x)$ est la différentielle de $F(x)$.

On montrera plus bas (chap. 6, § 3) que *si $f(x)$ est continue sur $]a, b[$, alors elle admet une primitive sur $]a, b[$,* donc, une intégrale indéfinie.

Voici quelques propriétés de l'intégrale indéfinie, résultant immédiatement de la définition.

1° $d\int f(x)\,dx = f(x)\,dx$. En effet, $\int f(x)\,dx = F(x) + C$, d'où

$$d\int f(x)\,dx = d(F(x) + C) = dF(x) = F'(x)\,dx = f(x)\,dx.$$

2° $\int dF(x) = F(x) + C$, c'est-à-dire que $\int$ et $d$ se simplifient mais il faut ajouter une constante $C$ à $F(x)$. En effet, $\int dF(x) = \int F'(x)\,dx =$ (par définition) $= F(x) + C$.

3° $\int Af(x)\,dx = A\int f(x)\,dx + C$, où $A$ et $C$ sont des constantes.

4° $\int [f(x) + g(x)]\,dx = \int f(x)\,dx + \int g(x)\,dx + C$, où $C$ est une constante.

En effet

$$\left(\int f(x)\,dx+\int g(x)\,dx\right)'=\left(\int f(x)\,dx\right)'+\left(\int g(x)\,dx\right)'=$$
$$=(\text{par définition})=f(x)+g(x).$$

D'autre part,

$$\left(\int [f(x)+g(x)]\,dx\right)'=(\text{par définition})=f(x)+g(x).$$

Donc, les fonctions $\int f\,dx+\int g\,dx$ et $\int [f+g]\,dx$ sont les primitives d'une même fonction $f+g$. Par suite, elles diffèrent entre elles d'une constante $C$.

5° Si $F(x)$ est une primitive de $f(x)$, alors

$$\int f(ax+b)\,dx=\frac{1}{a}F(ax+b)+C.$$

En effet,

$$\left[\frac{1}{a}F(ax+b)\right]'=\frac{1}{a}\cdot aF'(ax+b)=f(ax+b).$$

Formons le tableau suivant en se servant des principales formules du calcul différentiel.

1. $\int 0\cdot dx=C.$

2. $\int x^{\alpha}\,dx=\frac{x^{\alpha+1}}{\alpha+1}+C,\quad \forall\alpha\neq-1.$

3. $\int x^{-1}\,dx=\int\frac{dx}{x}=\ln|x|+C$, sur un intervalle ne contenant pas $x=0$.

4. $\int a^x\,dx=\frac{a^x}{\ln a}+C\quad (0<a,\ a\neq 1),\quad \int e^x\,dx=e^x+C.$

5. $\int \sin x\,dx=-\cos x+C,\quad \int\cos x\,dx=\sin x+C.$

6. $\int\frac{dx}{\cos^2 x}=\operatorname{tg} x+C,\quad \int\frac{dx}{\sin^2 x}=-\operatorname{cotg} x+C$ sur un intervalle où la fonction à intégrer est continue.

7. $\int\frac{dx}{\sqrt{1-x^2}}=\begin{cases}\operatorname{Arcsin} x+C,\\ -\operatorname{Arccos} x+C\end{cases}\quad(-1<x<1).$

8. $\int\frac{dx}{1+x^2}=\begin{cases}\operatorname{Arctg} x+C,\\ -\operatorname{Arccotg} x+C.\end{cases}$

9. $\int \operatorname{sh} x\,dx=\operatorname{ch} x+C,\quad \int\operatorname{ch} x\,dx=\operatorname{sh} x+C.$

10. $\int \frac{dx}{\operatorname{ch}^2 x} = \operatorname{th} x + C, \quad \int \frac{dx}{\operatorname{sh}^2 x} = -\coth x + C \quad (x \neq 0).$

11. $\int \frac{dx}{\sqrt{x^2+1}} = \ln |x + \sqrt{x^2+1}| + C = \operatorname{Arg\,sh} x + C,$

$\int \frac{dx}{\sqrt{x^2-1}} = \ln |x + \sqrt{x^2-1}| + C = \operatorname{Arg\,ch} x + C \quad (|x| > 1).$

12. $\int \frac{dx}{1-x^2} = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| + C \quad (|x| \neq 1).$

Prouvons la formule 3. Comme $|x|' = \operatorname{sgn} x$ et $x \operatorname{sgn} x = |x|$ pour $x \neq 0$, on a

$$(\ln |x| + C)' = \frac{1}{|x|} (|x|)' = \frac{\operatorname{sgn} x}{|x|} = \frac{1}{x} \quad (x \neq 0),$$

ce qu'on voulait.

Prouvons maintenant la formule 11:

$$(\ln |x + \sqrt{x^2+1}| + C)' = \frac{\operatorname{sgn}(x + \sqrt{x^2+1})}{|x + \sqrt{x^2+1}|} \left(1 + \frac{x}{\sqrt{x^2+1}}\right) =$$
$$= \frac{(x + \sqrt{x^2+1}) \operatorname{sgn}(x + \sqrt{x^2+1})}{\sqrt{x^2+1}\, |x + \sqrt{x^2+1}|} = \frac{1}{\sqrt{x^2+1}},$$

c.q.f.d.

Par ailleurs

$$\int \frac{dx}{\sqrt{x^2+1}} = \operatorname{Arg\,sh} x + C,$$

donc, en vertu du théorème 2,

$$\operatorname{Arg\,sh} x = \ln |x + \sqrt{x^2+1}| + C.$$

Or, $\operatorname{Arg\,sh} 0 = 0$, donc

$$\ln |x + \sqrt{x^2+1}| = \operatorname{Arg\,sh} x \quad \text{(cf. chap. 4, § 6, n° 9).}$$

En se servant de la propriété 5°, on peut obtenir les intégrales de fonctions plus compliquées. Exemple:

$$\int \sin(ax+b)\, dx = -\frac{1}{a} \cos(ax+b) + C.$$

Signalons que si la dérivation d'une fonction élémentaire nous donne nécessairement une fonction élémentaire, l'intégration d'une fonction élémentaire peut nous conduire à une fonction non élémentaire, c'est-à-dire une fonction que l'on ne peut obtenir ni par superposition de fonctions élémentaires ni par un nombre fini d'opérations arithmétiques sur des fonctions élémentaires.

On a prouvé par exemple que les intégrales suivantes n'étaient pas des fonctions élémentaires:

$\int e^{-x^2}\,dx$ intégrale de Poisson,

$\int \cos x^2\,dx$, $\int \sin x^2\,dx$ intégrales de Fresnel,

$\int \frac{dx}{\ln x}$ logarithme intégral,

$\int \frac{\cos x}{x}\,dx$ cosinus intégral,

$\int \frac{\sin x}{x}\,dx$ sinus intégral.

Ces intégrales existent mais ne sont pas des fonctions élémentaires. Il existe d'autres méthodes pour les calculer. Le sinus intégral, par exemple, peut être représenté par une série entière illimitée (cf. chap. 4, § 16)

$$\int \frac{\sin x}{x}\,dx = C + x - \frac{x^3}{3\cdot 3!} + \frac{x^5}{5\cdot 5!} + \ldots$$

## § 2. Méthodes d'intégration

La formule d'*intégration par changement de variable*

$$\int f(x)\,dx = \int f(\varphi(t))\,\varphi'(t)\,dt + C = \int f(\varphi(t))\,d\varphi(t) + C \qquad (1)$$

est essentielle dans le calcul intégral. Dans cette formule on admet que $x = \varphi(t)$ est une fonction continûment dérivable (c'est-à-dire possédant une dérivée continue) sur un intervalle de variation de $t$, $f(x)$, une fonction continue sur l'intervalle correspondant de l'axe $Ox$. La première égalité (1) affirme que le premier membre est identiquement égal au second si (après intégration! ) on fait la substitution $x = \varphi(t)$ et que l'on choisisse convenablement la constante $C$. Prouvons ceci. Au premier membre de (1) figure une fonction qui est une primitive de $f(x)$. Sa dérivée par rapport à $t$ est

$$\frac{d}{dt}\int f(x)\,dx = \frac{d}{dx}\left(\int f(x)\,dx\right)\frac{dx}{dt} = f(\varphi(t))\,\varphi'(t).$$

Donc, si l'on fait la substitution $x = \varphi(t)$ dans cette fonction, on obtient une primitive de la fonction $f(\varphi(t))\,\varphi'(t)$. L'intégrale du second membre est par définition une primitive de $f(\varphi(t))\,\varphi'(t)$. Or, deux primitives d'une fonction diffèrent entre elles d'une constante $C$. C'est ce qui est exprimé dans la première égalité (1). Quant à la deuxième, elle revêt un caractère purement formel: on conviendra d'écrire simplement

$$\int F(t)\,\varphi'(t)\,dt = \int F(t)\,d\varphi(t). \qquad (2)$$

Exemple:

$$\int e^{x^2}x\,dx=\frac{1}{2}\int e^{x^2}2x\,dx+C=\frac{1}{2}\int e^{x^2}\,dx^2+C=$$
$$=\frac{1}{2}\int e^u\,du+C_1=\frac{1}{2}e^u+C_2=\frac{1}{2}e^{x^2}+C_2\quad (u=x^2). \qquad (3)$$

La première égalité résulte de 3° du § 1, la seconde, de (2), la troisième de (1) (avec une autre constante), la quatrième, de la formule 4 du tableau. L'omission de la constante $C$ dans les étapes intermédiaires simplifie les calculs. Ainsi, dans l'exemple précédent:

$$\int e^{x^2}x\,dx=\frac{1}{2}\int e^{x^2}2x\,dx=\frac{1}{2}\int e^{x^2}\,dx^2=\frac{1}{2}e^{x^2}+C.$$

Voici un autre exemple: $I=\int\sqrt{a^2-x^2}\,dx$, $a>0$. Cette intégrale ne figure pas au tableau. En posant $x=a\sin t$, on obtient $\sqrt{a^2-x^2}=a\sqrt{1-\sin^2 t}=a\cos t$ et $dx=a\cos t\,dt$. Donc,

$$I=\int a\cos t\,a\cos t\,dt=a^2\int\cos^2 t\,dt=a^2\int\frac{1+\cos 2t}{2}\,dt=$$
$$=\frac{a^2t}{2}+\frac{a^2}{4}\sin 2t+C.$$

Or, $t=\operatorname{Arcsin}\frac{x}{a}$, donc

$$I=\frac{a^2}{2}\operatorname{Arcsin}\frac{x}{a}+\frac{a^2}{2}\sin t\cos t+C=\frac{a^2}{2}\operatorname{Arcsin}\frac{x}{a}+$$
$$+\frac{a^2}{2}\frac{x}{a}\sqrt{1-\left(\frac{x}{a}\right)^2}+C=\frac{a^2}{2}\operatorname{Arcsin}\frac{x}{a}+\frac{x}{2}\sqrt{a^2-x^2}+C.$$

En définitive

$$\int\sqrt{a^2-x^2}\,dx=\frac{x}{2}\sqrt{a^2-x^2}+\frac{a^2}{2}\operatorname{Arcsin}\frac{x}{a}+C.$$

Citons d'autres exemples qui nous seront utiles dans l'intégration des fonctions rationnelles:

$$\int\frac{dx}{(x-a)^m}=\int\frac{d\,(x-a)}{(x-a)^m}=\frac{1}{(x-a)^{m-1}(1-m)}+C\quad (m\neq 1); \qquad (4)$$

$$\int\frac{dx}{x-a}=\int\frac{d\,(x-a)}{x-a}=\ln|x-a|+C; \qquad (5)$$

$$\int\frac{dx}{x^2+a^2}=\frac{1}{a}\int\frac{d\,(x/a)}{1+(x/a)^2}=\frac{1}{a}\operatorname{Arctg}\frac{x}{a}+C;$$

$$\int\frac{dx}{x^2-a^2}=\frac{1}{2a}\int\left(\frac{1}{x-a}-\frac{1}{x+a}\right)dx=$$
$$=\frac{1}{2a}(\ln|x-a|-\ln|x+a|)+C=\frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right|+C;$$

$$\int \frac{dx}{x^2+px+q} = \int \frac{dx}{(x+(p/2))^2} =$$
$$= \int \frac{d\,(x+(p/2))}{(x+(p/2))^2} = -\frac{1}{x+(p/2)} + C \quad \left(q-\frac{p^2}{4}=0\right);$$

$$\int \frac{dx}{x^2+px+q} = \int \frac{dx}{(x+(p/2))^2+(q-(p^2/4))} =$$
$$= \int \frac{d\,(x+(p/2))}{(x+(p/2))^2+a^2} = \frac{1}{a}\int \frac{d\left(\dfrac{x+(p/2)}{a}\right)}{1+\left(\dfrac{x+(p/2)}{a}\right)^2} =$$
$$= \frac{1}{a}\,\mathrm{Arctg}\,\frac{x+(p/2)}{a} + C \quad \left(q-\frac{p^2}{4}=a^2,\ a>0\right); \tag{6}$$

$$\int \frac{dx}{x^2+px+q} = \int \frac{d\,(x+(p/2))}{(x+(p/2))^2-a^2} =$$
$$= \frac{1}{2a}\ln\left|\frac{x+(p/2)-a}{x+(p/2)+a}\right| + C \quad \left(q-\frac{p^2}{4}=-a^2,\ a>0\right);$$

$$\int \frac{(2x+p)\,dx}{x^2+px+q} = \int \frac{d\,(x^2+px+q)}{x^2+px+q} = \ln|x^2+px+q| + C;$$

$$\int \frac{Ax+B}{x^2+px+q}\,dx = \frac{A}{2}\int \frac{2x+(2B/A)}{x^2+px+q}\,dx = \frac{A}{2}\int \frac{(2x+p)\,dx}{x^2+px+q} +$$
$$+\frac{A}{2}\int \frac{(2B/A)-p}{x^2+px+q}\,dx = \frac{A}{2}\ln|x^2+px+q| + D\int \frac{dx}{x^2+px+q}, \tag{7}$$

$\left(A\neq 0,\ D=\frac{A}{2}\left(\frac{2B}{A}-p\right)\right)$ (voir (6)).

Le fait important ici, c'est que les intégrales de type (4) à (7), où $a$, $A$, $B$, $p$ et $q$ sont des constantes, sont des fonctions élémentaires (rationnelles, ln et Arctg).

Passons maintenant à la *formule d'intégration par parties*:

$$\int uv'\,dx = uv - \int vu'\,dx + C \tag{8}$$

ou ce qui revient au même

$$\int u\,dv = uv - \int v\,du + C.$$

On omet généralement la constante $C$, car le second membre de (8) est une intégrale indéfinie.

Dans cette formule, on admet que $u(x)$ et $v(x)$ sont des fonctions continûment dérivables. La validité de la formule (8) résulte du fait que les dérivées du premier et du second membre sont égales:

$$uv' = (uv)' - vu'.$$

La formule (8) ramène le calcul de l'intégrale $\int u\,dv$ à celui de $\int v\,du$. La méthode d'intégration par la formule (8) s'appelle *méthode d'intégration par parties.*

EXEMPLE 1. Calculer $\int x \ln x\,dx$. Posons

$$u(x) = \ln x, \quad du = \frac{dx}{x},$$
$$x\,dx = dv, \quad v = \int dv = \int x\,dx = \frac{x^2}{2}.$$

Alors

$$\int x \ln x\,dx = \frac{x^2}{2}\ln x - \int \frac{x^2}{2}\frac{dx}{x} = \frac{x^2}{2}\ln x - \frac{x^2}{4} + C.$$

EXEMPLE 2. Calculer les intégrales $I = \int e^{ax} \sin bx\,dx$, $I_1 = \int e^{ax} \cos bx\,dx$, où $a$ et $b$ sont des constantes. L'intégrant peut être mis de deux manières sous forme d'un produit: $u = e^{ax}$, $dv = \sin bx\,dx$ ou $u = \sin bx$, $dv = e^{ax}\,dx$.

Supposons que

$$u = e^{ax}, \quad du = ae^{ax}\,dx,$$
$$\sin bx\,dx = dv, \quad v = -\frac{\cos bx}{b}.$$

Une intégration par parties nous donne

$$I = -\frac{1}{b}e^{ax}\cos bx + \frac{a}{b}\int e^{ax}\cos bx\,dx = -\frac{1}{b}e^{ax}\cos bx + \frac{a}{b}I_1. \quad (9)$$

Intégrons $I_1$ par parties en posant $u = e^{ax}$, $dv = \cos bx\,dx$. On obtient

$$I_1 = \frac{1}{b}e^{ax}\sin bx - \frac{a}{b}I. \quad (10)$$

De (9) et de (10), on déduit le système

$$\begin{cases} I - \frac{a}{b}I_1 = -\frac{1}{b}e^{ax}\cos bx, \\ \frac{a}{b}I + I_1 = \frac{1}{b}e^{ax}\sin bx, \end{cases}$$

dont la résolution nous donne

$$I = \frac{a\sin bx - b\cos bx}{a^2+b^2}e^{ax} + C, \quad I_1 = \frac{b\sin bx + a\cos bx}{a^2+b^2}e^{ax} + C.$$

EXEMPLE 3. Calculer l'intégrale $I = \int \operatorname{Arcsin} x\,dx$.

En posant $u = \text{Arcsin}\, x$, $dv = dx$, on obtient

$$I = x \text{ Arcsin}\, x - \int \frac{x\, dx}{\sqrt{1-x^2}} = x \text{ Arcsin}\, x + \sqrt{1-x^2} + C.$$

EXEMPLE 4. Citons un exemple qui nous servira dans l'intégration des fonctions rationnelles. Soit à calculer $I = \int \frac{dx}{(x^2+a^2)^{k-1}}$, où $k$ est un entier naturel $>1$ et $a > 0$. On a

$$\int \frac{dx}{(x^2+a^2)^{k-1}} = a^2 \int \frac{dx}{(x^2+a^2)^k} + \frac{1}{2} \int \frac{x \cdot 2x\, dx}{(x^2+a^2)^k}.$$

En posant $u = x$ et $dv = \frac{2x\, dx}{(x^2+a^2)^k}$, on obtient

$$I = a^2 \int \frac{dx}{(x^2+a^2)^k} + \frac{1}{2} \left\{ \frac{x}{(1-k)(x^2+a^2)^{k-1}} - \frac{1}{1-k} \int \frac{dx}{(x^2+a^2)^{k-1}} \right\},$$

d'où

$$a^2 \int \frac{dx}{(x^2+a^2)^k} = \frac{x}{2(k-1)(x^2+a^2)^{k-1}} + \frac{2k-3}{2(k-1)} \int \frac{dx}{(x^2+a^2)^{k-1}}.$$

On peut maintenant (si $k > 2$) reprendre cette procédure pour l'intégrale du second membre et abaisser ainsi d'une unité l'exposant du dénominateur de l'intégrant. On arrive en définitive à une intégrale de la fonction $(x^2 + a^2)^{-1}$ qui, on le sait, est égale à $\frac{1}{a}$ Arctg $\frac{x}{a} + C$.

Donc, pour $q - (p^2/4) = a^2 > 0$ et $k$ entier naturel, l'intégrale

$$\int \frac{dx}{(x^2+px+q)^k} = \int \frac{du}{(u^2+a^2)^k} + C \quad \left(u = x + \frac{p}{2}\right) \qquad (11)$$

se compose de fonctions élémentaires.

EXEMPLE 5. Calculer les intégrales

$$\int P_n(x) \begin{Bmatrix} e^{bx} \\ \cos bx \\ \sin bx \end{Bmatrix} dx,$$

où $P_n(x) = a_n x^n + \ldots + a_1 x + a_0$ est un polynôme de degré $n$.

On calcule ces intégrales par $n$ intégrations par parties en posant successivement $u = P_n(x)$, $u = P_n'(x)$, ... Les intégrales obtenues se simplifieront, car la dérivée d'un polynôme algébrique $P_n(x)$ de degré $n$ est un polynôme algébrique de degré $n - 1$.

Les intégrales de ce type se calculent par la *méthode des coefficients indéterminés.*

Par exemple, $\int P_n(x)\, e^{bx}\, dx$ est une fonction de la forme $Q_n(x)\, e^{bx} + C$, où $Q_n(x) = b_n x^n + \ldots + b_1 x + b_0$. Les coeffi-

cients inconnus $b_0, \ldots, b_n$ se déterminent à partir de la condition

$$(Q_n(x)\, e^{bx} + C)' = P_n(x)\, e^{bx} \quad \text{ou} \quad Q_n'(x) + bQ_n(x) = P_n(x).$$

L'identification des coefficients nous donne $b_0, \ldots, b_n$. On s'est servi du fait que deux polynômes sont égaux si et seulement si leurs coefficients en les puissances respectives de $x$ sont égaux (cf. chap. 4, § 14, théorème 2).

Illustrons ceci sur un exemple suivant:

$$\int (x^2+1)\, e^x\, dx = (ax^2+bx+c)\, e^x + C.$$

On a ici

$$P_2(x) = x^2+1,\ Q_2(x) = ax^2+bx+c,$$

les coefficients $a$, $b$, $c$ étant à déterminer. On a

$$(Q_2(x)\, e^x)' = [ax^2 + (2a+b)\, x + b + c]\, e^x = (x^2+1)\, e^x,$$

d'où $ax^2 + (2a+b)\, x + b + c = x^2 + 1$. Donc $a = 1$, $2a + b = 0$, $b + c = 1$. On obtient en définitive

$$\int (x^2+1)\, e^x\, dx = (x^2-2x+3)\, e^x + C.$$

## § 3. Nombres complexes

On appelle *nombre complexe* un nombre de la forme

$$z = a + bi = a + ib,$$

où $a$ et $b$ sont des nombres réels, $i$ un symbole spécial. Les nombres complexes sont justiciables de la notion d'égalité et des opérations arithmétiques. Soient deux nombres complexes $z_1 = a_1 + ib_1$, $z_2 = a_2 + ib_2$.

1) $z_1 = z_2$ si et seulement si $a_1 = a_2$ et $b_1 = b_2$; $a + 0i = a$, $0 + bi = bi$, $1 \cdot i = i$.

2) $z_1 \pm z_2 = (a_1 \pm a_2) + i\,(b_1 \pm b_2)$.

3) $z_1 \cdot z_2 = (a_1a_2 - b_1b_2) + i\,(b_1a_2 + a_1b_2)$.

4) $\dfrac{z_1}{z_2} = \dfrac{a_1a_2+b_1b_2}{a_2^2+b_2^2} + i\,\dfrac{b_1a_2-a_1b_2}{a_2^2+b_2^2} \quad (a_2^2+b_2^2 \neq 0)$.

De 1) et 3) il s'ensuit que

$$i^2 = -1.$$

Donc, les opérations d'addition et de multiplication introduites sont commutatives ($z_1 + z_2 = z_2 + z_1$, $z_1z_2 = z_2z_1$), associatives ($(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$, $(z_1z_2)\, z_3 = z_1\,(z_2z_3)$), distributives ($(z_1 + z_2)\, z_3 = z_1z_2 + z_2z_3$).

De la propriété $a + 0i = a$ il résulte que l'ensemble des nom-

bres complexes contient celui des réels. Il est immédiat de voir en outre que l'application des opérations arithmétiques 2), 3) et 4) aux nombres $z_1 = a_1 + 0i$, $z_2 = a_2 + 0i$ nous donne respectivement
$a_1 \pm a_2 + 0i = a_1 \pm a_2$, $a_1a_2 + 0i = a_1a_2$, $\frac{a_1}{a_2} + 0i = \frac{a_1}{a_2} (a_2 \neq 0)$.

Le nombre $\bar{z} = a - ib$ est le *conjugué complexe de* $z$. Le nombre réel $|z| = \sqrt{a^2 + b^2}$ s'appelle *module du nombre complexe* $z$. Il est évident que $z \cdot \bar{z} = |z|^2$.

Si l'on interprète le nombre complexe $z = a + ib$ comme un point $M(a, b)$ du plan $xOy$, alors $|z|$ est égal à la distance de $M$ à l'origine des coordonnées $O$ (fig. 74).

En introduisant les coordonnées polaires $(\rho, \varphi)$, on obtient

$$\begin{aligned} a &= \rho \cos \varphi = |z| \cos \varphi, \\ b &= \rho \sin \varphi = |z| \sin \varphi \qquad (1) \\ & (|z| > 0). \end{aligned}$$

Le nombre complexe $z$ s'écrit alors

$$z = \rho (\cos \varphi + i \sin \varphi), \qquad (2)$$

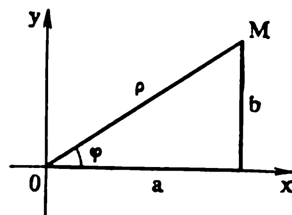


Fig. 74

où $\rho$ est le module du nombre $z$, $\varphi$, l'angle (en radians) du vecteur $\overrightarrow{OM}$ avec $Ox$. Cet angle s'appelle encore *argument principal du nombre complexe* $z$ et se note: $\varphi = \text{Arg } z$ $(0 \leqslant \varphi < 2\pi)$.

Il est évident que $\varphi = \text{Arg } z$ est une fonction de $z$, $z \neq 0$. On introduit encore la correspondance

$$\varphi = \arg z = \text{Arg } z + 2k\pi \quad (k = 0, \pm 1, \pm 2, \ldots)$$

qui donne toutes les valeurs de $\varphi$ vérifiant (1) pour $z \neq 0$ donné.

Le nombre $z = 0$ est le seul à ne pas avoir d'argument, par contre on peut le définir comme un nombre dont le module est nul ($|z| = 0$).

Arg $z$ s'appelle encore *argument sous la forme réduite*. Il y a parfois intérêt à admettre que Arg $z$ appartient à un autre intervalle semi-ouvert $[a, a + 2\pi[$, par exemple à $[-\pi, \pi[$.

Le nombre $a$ s'appelle *partie réelle* de $z$ et se note $a = \text{Re } z$, le nombre $b$, *partie imaginaire* de $z$ et se note $b = \text{Im } z$. Donc

$$z = \text{Re } z + i \text{ Im } z.$$

Par définition

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \qquad (-\infty < \varphi < \infty). \qquad (3)$$

Il est évident que $e^{i\varphi}$ est une fonction complexe (c'est-à-dire à valeurs complexes) de l'argument réel $\varphi$. La fonction $e^{i\varphi}$ est visiblement $2\pi$-périodique: $e^{i(\varphi+2\pi)} = e^{i\varphi}$.

Comme $|e^{i\varphi}| = \sqrt{\cos^2\varphi + \sin^2\varphi} = 1$, le point $e^{i\varphi}$ décrit continûment un cercle de rayon 1 centré en $z = 0$, lorsque $\varphi$ varie continûment sur l'intervalle semi-ouvert $[0, 2\pi[$.

On a les relations

$$e^{i(\varphi_1+\varphi_2)} = e^{i\varphi_1}e^{i\varphi_2}, \quad e^{-i\varphi} = \frac{1}{e^{i\varphi}}. \tag{4}$$

En effet

$$\begin{aligned} e^{i(\varphi_1+\varphi_2)} &= \cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2) = \\ &= (\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + \\ &+ i(\sin\varphi_1\cos\varphi_2 + \sin\varphi_2\cos\varphi_1) = \\ &= (\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_1) = e^{i\varphi_1}e^{i\varphi_2}, \end{aligned}$$

$$\frac{1}{e^{i\varphi}} = \frac{1}{\cos\varphi + i\sin\varphi} = \cos\varphi - i\sin\varphi = \cos(-\varphi) + i\sin(-\varphi) = e^{-i\varphi}.$$

La fonction $e^z$, où $z = x + iy$ est une variable complexe arbitraire. se définit par

$$e^z = e^x e^{iy}, \quad z \neq 0.$$

En vertu de (3) on a

$$e^z = e^x(\cos y + i\sin y). \tag{5}$$

En tenant compte de (2) et de (3) on peut représenter tout nombre complexe $z$ par

$$z = \rho e^{i\varphi} \quad (\rho \geqslant 0), \tag{6}$$

où $\rho = |z|$ est unique et

$$\varphi = \arg z = \operatorname{Arg} z + 2k\pi \quad (k = 0, \pm1, \pm2, \ldots).$$

Les expressions (2) et (6) s'appellent respectivement *formes trigonométrique* et *exponentielle du nombre complexe z*.

Voici quelques exemples de nombres complexes écrits sous la forme exponentielle (on admet que $\varphi = \operatorname{Arg} z$):

$$1 + i = \sqrt{2}\left(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right) = \sqrt{2}\, e^{i\pi/4},$$

$$i = 0 + 1\cdot i = e^{i\pi/2}, \quad 1 = e^{0i}, \quad -1 = e^{\pi i}.$$

Des relations (3) et (4) on déduit immédiatement la *formule de Moivre*

$$(\cos\varphi + i\sin\varphi)^n = e^{in\varphi} = \cos n\varphi + i\sin n\varphi. \tag{7}$$

On a aussi

$$z_1 z_2 = |z_1|\,|z_2|\, e^{i(\varphi_1+\varphi_2)},$$

c'est-à-dire que lorsqu'on multiplie des nombres complexes, leurs modules se multiplient et leurs arguments s'ajoutent.

La conjugaison des nombres complexes jouit des propriétés suivantes:

$$\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2} \quad (z_2 \neq 0). \tag{8}$$

En effet

$$\overline{(a_1 + b_1 i) \pm (a_2 + b_2 i)} = \overline{(a_1 \pm a_2) + (b_1 \pm b_2)\, i} =$$
$$= (a_1 \pm a_2) - i\,(b_1 \pm b_2) = (a_1 - b_1 i) \pm (a_2 - b_2 i) =$$
$$= \overline{(a_1 + b_1 i)} \pm \overline{(a_2 + b_2 i)};$$

d'autre part, comme

$$\overline{\rho e^{i\varphi}} = \overline{\rho\,(\cos\varphi + i\sin\varphi)} = \rho\,(\cos\varphi - i\sin\varphi) = \rho e^{-i\varphi},$$

il vient

$$\overline{z_1 z_2} = \overline{\rho_1 e^{i\varphi_1}\rho_2 e^{i\varphi_2}} = \overline{\rho_1\rho_2 e^{i(\varphi_1+\varphi_2)}} = \rho_1\rho_2 e^{-i(\varphi_1+\varphi_2)} =$$
$$= \rho_1 e^{-i\varphi_1}\rho_2 e^{-i\varphi_2} = \overline{\rho_1 e^{i\varphi_1}}\cdot\overline{\rho_2 e^{i\varphi_2}} = \bar{z}_1\cdot\bar{z}_2.$$

On procède de même pour le quotient.

On considère maintenant le problème suivant. Trouver les racines $n$-ièmes d'un nombre $a = \rho e^{i\theta}$ $(\rho > 0)$, c'est-à-dire tous les nombres $b = re^{i\varphi}$ tels que $b^n = a$. On a alors $r^n e^{in\varphi} = \rho e^{i\theta}$ $(r, \rho > 0)$, et par suite de l'unicité de la représentation d'un nombre complexe sous la forme exponentielle, il vient $\rho = r^n$, $n\varphi = \theta + 2k\pi$, $k = 0, \pm 1, \pm 2, \ldots$ De la première égalité il s'ensuit que $r = \sqrt[n]{\rho}$ ($r$ est la valeur arithmétique de la racine $n$-ième du nombre strictement positif $\rho$). On déduit de la deuxième égalité que $\varphi = \frac{\theta}{n} + \frac{2k\pi}{n}$ $(k = 0, \pm 1, \ldots)$.

Les valeurs de $\varphi$ qui nous donnent des racines $n$-ièmes de $a$ distinctes correspondent aux $n$ premières valeurs de $k$:

$$\varphi_k = \frac{\theta}{n} + \frac{2k\pi}{n} \quad (k = 0, 1, \ldots, n-1). \tag{9}$$

Pour les autres $k$ on obtient des valeurs de $\varphi$ qui diffèrent de l'une des valeurs (9) d'un multiple de $2\pi$.

Nous avons ainsi montré qu'un nombre complexe $a \neq 0$ possède $n$ (et seulement $n$) racines $n$-ièmes exprimées par la formule

$$\sqrt[n]{a} = \sqrt[n]{\rho e^{i\theta}} = \sqrt[n]{\rho}\, e^{i\varphi_k} \quad (k = 0, 1, \ldots, n-1),$$

où $\varphi_k$ sont définis par (9).

EXEMPLES:

1. $\sqrt[3]{1}=\sqrt[3]{e^{0i}}=e^{\left(\frac{0}{3}+\frac{2k\pi}{3}\right)i}$ $(k=0, 1, 2)$.

2. $\sqrt[3]{i}=\sqrt[3]{e^{i\frac{\pi}{2}}}=e^{\left(\frac{\pi}{6}+\frac{2k\pi}{3}\right)i}$ $(k=0, 1, 2)$.

3. $\sqrt[6]{1+i}=\sqrt[12]{2}\sqrt[6]{e^{i\frac{\pi}{4}}}=\sqrt[12]{2}e^{i\left(\frac{\pi}{24}+\frac{2k\pi}{6}\right)}$ $(k=0, 1, \ldots, 5)$.

## § 4. Théorie du polynôme de degré $n$

On appelle *polynôme de degré n* une fonction de la forme

$$Q_n(z)=a_0+a_1z+\ldots+a_nz^n=\sum_{k=0}^{n}a_kz^k, \tag{1}$$

où $a_k$ sont des coefficients réels ou complexes, $z$, une variable généralement complexe, qui peut prendre n'importe quelle valeur complexe $(z=x+iy)$ ou, dans le langage géométrique, être un point quelconque du plan complexe.

A tout point $z$ du plan complexe la formule (1) associe un nombre $Q_n(z)$ généralement complexe. On admettra dans la suite que $a_n\neq 0$. Si $Q_n(a)=0$, le nombre $a$ s'appelle *racine* ou *zéro* du polynôme $Q_n(z)$.

En raisonnant exactement comme au début du § 14 du chap. 4 pour le cas d'un polynôme d'une variable réelle, on démontre que quel que soit le nombre complexe $z_0$, le polynôme $Q_n(z)$ se développe en une série entière de $z-z_0$ et ce de façon unique, c'est-à-dire que

$$Q_n(z)=\sum_{k=0}^{n}b_k(z-z_0)^k,$$

où $b_k$ sont des coefficients généralement complexes. Il est évident que $Q_n(z_0)=b_0$. Il s'ensuit qu'une condition nécessaire et suffisante pour que $z_0$ soit racine de $Q_n$ est que le terme constant du développement de $Q_n$ en série entière de $z-z_0$ soit nul. Si $b_0=0$, on peut mettre $Q_n$ sous la forme

$$Q_n(z)=(z-z_0)\,Q_{n-1}(z),\quad \forall z, \tag{2}$$

où $Q_{n-1}$ est un polynôme de degré $n-1$. Réciproquement, si $Q_n(z)$ se met sous la forme (2), autrement dit si $Q_n(z)$ est divisible par $z-z_0$, alors $z_0$ est de toute évidence un zéro de $Q_n$.

Nous avons prouvé le

THÉORÈME DE BÉZOUT. *Pour qu'un polynôme $Q_n(z)$ possède une racine (complexe) $z_0$ il est nécessaire et suffisant qu'il soit divisible par $z-z_0$, c'est-à-dire qu'il se représente sous la forme* (2), *où $Q_{n-1}$ est un polynôme de degré $n-1$.*

Supposons que $z_0$ est un zéro de $Q_n$, donc que $Q_n$ se représente par (2). Si $Q_{n-1}(z_0) \neq 0$, le théorème de Bézout appliqué à $Q_{n-1}$ nous dit que $Q_{n-1}(z)$ n'est pas divisible par $z - z_0$. On dit alors que $z_0$ est une *racine* (ou un *zéro*) *simple* de $Q_n$. Supposons maintenant que $Q_{n-1}(z_0) = 0$. Le théorème de Bézout appliqué à $Q_{n-1}(z)$ nous dit que ce dernier est divisible par $z - z_0$. On obtient donc: $Q_n(z) = (z - z_0)^2 Q_{n-2}(z)$, où $Q_{n-2}(z)$ est un polynôme de degré $n - 2$. Si $Q_{n-2}(z_0) \neq 0$, on dit alors que $z_0$ est une *racine* (ou un *zéro*) *double* ou *de multiplicité* 2 de $Q_n$. D'une façon plus générale, si

$$Q_n(z) = (z - z_0)^s Q_{n-s}(z), \qquad Q_{n-s}(z_0) \neq 0, \qquad s \leqslant n,$$

on dit alors que $z_0$ *est une racine* (*un zéro*) *de multiplicité s*.

On a le théorème d'existence d'une racine complexe d'un polynôme.

Théorème fondamental. *Tout polynôme de degré n possède une racine complexe au moins.*

Nous glisserons sur la démonstration de ce théorème.

Voici un corollaire important.

Corollaire. *Un polynôme $Q_n$ de degré n, dont le coefficient dominant $a_n \neq 0$, possède n racines complexes compte tenu de leur multiplicité, autrement dit*

$$Q_n(z) = a_n (z - z_1)^{p_1} (z - z_2)^{p_2} \dots (z - z_l)^{p_l}, \tag{3}$$
$$p_1 + p_2 + \dots + p_l = n,$$

*où $z_1, \dots, z_l$ sont des racines distinctes de multiplicités respectives $p_1, \dots, p_l$.*

Démonstration. D'après le théorème fondamental, le polynôme $Q_n$ admet au moins une racine. Soient $z_1$ cette racine et $p_1$ sa multiplicité. On a

$$Q_n(z) = (z - z_1)^{p_1} Q_{n-p_1}(z), \quad Q_{n-p_1}(z_1) \neq 0.$$

Si $n - p_1 = 0$, c'est-à-dire que $p_1 = n$, alors nécessairement $Q_{n-p_1}(z) = a_n$ et le théorème est prouvé. Dans ce cas, $Q_n(z) = a_n (z - z_1)^n$.

Si $p_1 < n$, alors $Q_{n-p_1}(z)$ est un polynôme de degré $n - p_1$ non divisible par $z - z_1$, dont le coefficient dominant est différent de zéro. On peut lui appliquer le théorème fondamental qui dit que $Q_{n-p_1}(z)$ possède une racine complexe. Désignons cette racine par $z_2$ et sa multiplicité par $p_2$. On obtient en définitive

$$Q_n(z) = (z - z_1)^{p_1} (z - z_2)^{p_2} Q_{n-p_1-p_2}(z)$$
$$(Q_{n-p_1-p_2}(z_j) \neq 0, \quad j = 1, 2).$$

Si $n - p_1 - p_2 = 0$, alors $Q_{n-p_1-p_2}(z) = a_n$. Sinon, on poursuit le processus. Celui-ci s'achève au bout d'un nombre fini ($\leqslant n$) de pas et l'on obtient la formule (3). Si dans le second membre de (3) on substitue à $z$ un nombre différent de $z_1, \dots, z_l$, on constate qu'il ne s'annule pas. Ceci montre qu'il n'existe pas de racines autres que celles trouvées et la représentation (3) est unique.

## § 5. Polynôme réel de degré $n$

On dit qu'un polynôme

$$Q_n(z)=\sum_{k=0}^{n} a_k z^k \quad (a_n \neq 0) \tag{1}$$

est *réel* si ses coefficients $a_k$ sont réels.

LEMME. *Pour tout polynôme réel $Q_n(z)$ on a*

$$Q_n(\bar{z})=\overline{Q_n(z)}, \quad \forall z.$$

DEMONSTRATION. Pour prouver ce lemme on se servira des égalités (8) du § 3 et du fait que $a_k = \bar{a}_k$ si $a_k$ est réel.

On a

$$Q_n(\bar{z})=\sum_{k=0}^{n} a_k \bar{z}^k=\sum_{k=0}^{n} \bar{a}_k \bar{z}^k=\sum_{k=0}^{n} \overline{a_k z^k}=\overline{\sum_{k=0}^{n} a_k z^k}=\overline{Q_n(z)}, \tag{2}$$

c.q.f.d.

THEOREME. *Si $z_0 = \alpha + i\beta$ $(\beta \neq 0)$ est une racine complexe de multiplicité $\nu$ d'un polynôme réel $Q_n$, alors il en est de même de $\bar{z}_0 = \alpha - i\beta$ et*

$$Q_n(z) = [(z-\alpha)^2 + \beta^2]^\nu Q_{n-2\nu}(z), \tag{3}$$

*où $Q_{n-2\nu}(z)$ est un polynôme réel de degré $n - 2\nu$ ne s'annulant pas pour $z = z_0$ et $z = \bar{z}_0$.*

DEMONSTRATION. Supposons que $z_0 = \alpha + i\beta$ $(\beta \neq 0)$ est une racine de $Q_n$. Alors il en est de même de $\bar{z}_0 = \alpha - i\beta$, car en vertu de (2) $Q_n(\bar{z}_0) = \overline{Q_n(z_0)} = \bar{0} = 0$. Les nombres $z_0$ et $\bar{z}_0$ ne sont pas égaux et $Q_n(z)$ est divisible par

$$(z-\alpha-i\beta)(z-\alpha+i\beta) = (z-\alpha)^2 + \beta^2. \tag{4}$$

Donc

$$Q_n(z) = [(z-\alpha)^2 + \beta^2]\, Q_{n-2}(z),$$

où $Q_{n-2}(z)$ est un polynôme de degré $n - 2$ qui est évidemment réel. En effet, le quotient de deux polynômes réels est un polynôme réel.

Si $z_0$ est une racine de $Q_n$ de multiplicité $\nu$, et $\nu > 1$, alors $z_0$ est une racine de $Q_{n-2}$ de multiplicité $\nu - 1$. Donc on peut, en reprenant les raisonnements précédents, représenter $Q_{n-2}(z)$ par le produit de (4) par un polynôme réel $Q_{n-4}$ de degré $n - 4$. En effectuant cette opération $\nu$ fois on obtient la représentation (3), où $Q_{n-2\nu}(z)$ est un polynôme réel de degré $n - 2\nu$ ne s'annulant pas en $z_0$. Alors $Q_{n-2\nu}(\bar{z}_0) \neq 0$. En effet, si $\bar{z}_0$ était racine de $Q_{n-2\nu}$, il en serait de même de $z_0$.

THÉORÈME 2. *Tout polynôme réel $Q_n(z)$, dont le coefficient dominant $a_n \neq 0$, peut être représenté par un produit de la forme*

$$Q_n(z) = a_n (z - c_1)^{\mu_1} \ldots (z - c_r)^{\mu_r} [(z - \alpha_1)^2 + \beta_1^2]^{\nu_1} \ldots$$
$$\ldots [(z - \alpha_s)^2 + \beta_s^2]^{\nu_s} = a_n \prod_{j=1}^{r} (z - c_j)^{\mu_j} \prod_{j=1}^{s} [(z - \alpha_j)^2 + \beta_j^2]^{\nu_j}, \quad (5)$$

*où $\beta_j > 0$, $\mu_1 + \ldots + \mu_r + 2(\nu_1 + \ldots + \nu_s) = n$, $c_1, \ldots, c_r$ sont les racines réelles de $Q_n$ de multiplicités respectives $\mu_1, \ldots, \mu_r$, et $\alpha_1 \pm \beta_1 i, \ldots, \alpha_s \pm \beta_s i$ les racines conjuguées complexes de $Q_n$ de multiplicités respectives $\nu_1, \ldots, \nu_s$.*

REMARQUE. Les polynômes réels du second degré de (5) peuvent être mis sous la forme:

$$(z - \alpha_j)^2 + \beta_j^2 = z^2 - 2\alpha_j z + (\alpha_j^2 + \beta_j^2) = z^2 + p_j z + q_j,$$
$$p_j = -2\alpha_j, \quad q_j = \alpha_j^2 + \beta_j^2.$$

Donc, la formule (5) peut s'écrire encore

$$Q_n(z) = a_n \prod_{j=1}^{r} (z - c_j)^{\mu_j} \prod_{j=1}^{s} (z^2 + p_j z + q_j)^{\nu_j}, \quad (5')$$

où $z^2 + p_j z + q_j$ sont des polynômes réels du second degré de racines complexes $\alpha_j \pm i\beta_j$ ($\beta_j > 0$, $p_j^2 - 4q_j = -4\beta_j^2 < 0$).

DÉMONSTRATION. D'après la formule (3) du § 4

$$Q_n(z) = \prod_{j=1}^{r} (z - c_j)^{\mu_j} Q_m(z),$$

où $Q_m(z)$ est un polynôme réel de degré $m = n - \mu_1 - \ldots - \mu_r$. Si $m = 0$, alors de toute évidence $Q_m(z) = a_n$; dans le cas général, on applique successivement le théorème 1 aux racines complexes de $Q_m$.

Signalons que le théorème fondamental n'affirme que l'existence d'une racine (généralement complexe) d'un polynôme de degré $n$, il n'indique aucune méthode pour trouver cette racine dans le cas général. Ce théorème se démontre par les méthodes de l'analyse mathématique. Nous passerons sur cette démonstration, car elle est intimement liée à la théorie des fonctions d'une variable complexe.

Il existe des formules pour la résolution des équations générales du second, du troisième et du quatrième degré. Abel [1]) a démontré qu'il n'existait pas de formule générale pour la résolution d'une équation de degré $n > 4$, c'est-à-dire que les racines de l'équation $a_n x^n + \ldots + a_1 x + a_0 = 0$ $(a_n \neq 0)$ ne s'expriment pas en fonction des coefficients $a_k$.

[1]) Niels Henrik Abel, illustre mathématicien norvégien (1802-1829).

## § 6. Intégration d'expressions rationnelles

On appelle *fonction* ou *fraction rationnelle* le quotient de deux polynômes algébriques

$$f(x) = \frac{P_m(x)}{Q_n(x)}, \tag{1}$$

$$P_m(x) = b_0 + b_1 x + \ldots + b_m x^m,$$
$$Q_n(x) = a_0 + a_1 x + \ldots + a_n x^n,$$

$b_m$, $a_n \neq 0$, $m \geqslant 0$, $n \geqslant 1$.

On conviendra que *la fonction rationnelle $f$ est réelle*, c'est-à-dire que $P_m$ et $Q_n$ sont des polynômes réels. On admettra de plus que $x$ est une variable réelle.

Les fractions rationnelles de la forme

$$\left\{ \begin{array}{l} \dfrac{A}{x-a}, \quad \dfrac{A}{(x-a)^k} \quad (k \geqslant 2), \quad \dfrac{Ax+B}{x^2+px+q}, \\ \qquad \dfrac{Ax+B}{(x^2+px+q)^k} \quad (k \geqslant 2), \end{array} \right. \tag{2}$$

où $A$, $B$, $a$, $p$, $q$ sont des nombres réels, $k$ un entier naturel, et le trinôme du second degré $x^2 + px + q$ ne possède pas de racines réelles, s'appellent *éléments* ou *fractions simples*.

Au § 2 on a montré comment on calcule les intégrales de fractions simples (voir (4), (5), (6), (7), (11) du § 2).

Soit à calculer l'intégrale indéfinie d'une fraction rationnelle $f(x)$ (voir (1)). Si $m \geqslant n$, une division ordinaire nous donne la partie entière de $f$:

$$f(x) = \text{un polynôme} + \frac{P_{m_1}(x)}{Q_n(x)} \quad (m_1 < n).$$

L'intégration du polynôme ne posant pas de difficultés, le problème se ramène à l'intégration d'une fraction rationnelle dont l'exposant du numérateur est inférieur à celui du dénominateur.

On admettra de ce fait que la fraction rationnelle $f(x)$ est *propre*, c'est-à-dire que l'exposant du numérateur est inférieur à celui du dénominateur $(m < n)$.

THÉORÈME. *Supposons que le dénominateur d'une fraction rationnelle réelle propre se représente par la formule* (5') *du* § 5:

$$Q_n(x) = a_n (x - c_1)^{\mu_1} \ldots (x - c_r)^{\mu_r} (x^2 + p_1 x + q_1)^{\nu_1} \ldots$$
$$\ldots (x^2 + p_s x + q_s)^{\nu_s}.$$

*Alors la fraction* (1) *se représente d'une façon unique par une somme de fractions simples*:

$$\frac{P_m(x)}{Q_n(x)}=\frac{A_{1,1}}{(x-c_1)^{\mu_1}}+\frac{A_{1,2}}{(x-c_1)^{\mu_1-1}}+\ldots+\frac{A_{1,\mu_1}}{x-c_1}+$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+\frac{A_{r,1}}{(x-c_r)^{\mu_r}}+\frac{A_{r,2}}{(x-c_r)^{\mu_r-1}}+\ldots+\frac{A_{r,\mu_r}}{x-c_r}+$$
$$+\frac{B_{1,1}x+C_{1,1}}{(x^2+p_1x+q_1)^{\nu_1}}+\frac{B_{1,2}x+C_{1,2}}{(x^2+p_1x+q_1)^{\nu_1-1}}+\ldots+\frac{B_{1,\nu_1}x+C_{1,\nu_1}}{x^2+p_1x+q_1}+$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$+\frac{B_{s,1}x+C_{s,1}}{(x^2+p_sx+q_s)^{\nu_s}}+\frac{B_{s,2}x+C_{s,2}}{(x^2+p_sx+q_s)^{\nu_s-1}}+\ldots+\frac{B_{s,\nu_s}x+C_{s,\nu_s}}{x^2+p_sx+q_s}, \qquad (3)$$

*où $A$, $B$ et $C$ (affectés de leurs indices respectifs) sont des constantes.*

Ce théorème affirme que pour toute fraction rationnelle réelle propre il existe des constantes $A$, $B$ et $C$ (affectées de leurs indices respectifs) telles que l'on ait (3) pour tous les $x$ sauf pour les valeurs $x = c_1, \ldots, c_r$ pour lesquelles les deux membres de (3) ne sont pas définis. Nous passerons sur la démonstration.

Illustrons ce théorème sur un exemple. Le théorème nous dit que

$$\frac{2x^3+x^2+x+2}{(x-1)^2(x^2+x+1)}=\frac{A_1}{x-1}+\frac{A_2}{(x-1)^2}+\frac{Mx+N}{x^2+x+1}, \qquad (4)$$

où $A_1$, $A_2$, $M$ et $N$ sont des constantes bien définies. Pour les trouver on réduit (4) au dénominateur commun, puis on égale les numérateurs du premier et du second membre:

$$2x^3+x^2+x+2=A_1(x-1)(x^2+x+1)+$$
$$+A_2(x^2+x+1)+(Mx+N)(x-1)^2. \qquad (5)$$

En chassant les parenthèses dans le second membre de (5) et en réduisant les termes semblables (voir chap. 4, § 14, théorème 2), on obtient:

$$\left\{\begin{aligned} 2&=A_1+M,\\ 1&=A_2+N-2M,\\ 1&=A_2+M-2N,\\ 2&=-A_1+A_2+N.\end{aligned}\right. \qquad (6)$$

On a un système de quatre équations linéaires à quatre inconnues $A_1$, $A_2$, $M$, $N$. D'après le théorème ci-dessus ce système admet

une solution unique: $A_1 = 1$, $A_2 = 2$, $M = N = 1$. Donc

$$\frac{2x^3+x^2+x+2}{(x-1)^2(x^2+x+1)} = \frac{1}{x-1} + \frac{2}{(x-1)^2} + \frac{x+1}{x^2+x+1}. \tag{7}$$

Et

$$\int \frac{2x^3+x^2+x+2}{(x-1)^2(x^2+x+1)}\,dx = \int \frac{dx}{x-1} + 2\int \frac{dx}{(x-1)^2} + \int \frac{x+1}{x^2+x+1}\,dx =$$

$$= \ln|x-1| - \frac{2}{x-1} + \ln\sqrt{x^2+x+1} + \frac{1}{\sqrt{3}}\,\mathrm{Arctg}\,\frac{2x+1}{\sqrt{3}} + C.$$

REMARQUE 1. L'égalité (5) est vérifiée pour tout $x \neq 1$. Elle est donc aussi vraie pour $x = 1$, car les fonctions du premier et du second membre sont des fonctions continues de $x$. En faisant $x = 1$, on obtient $6 = 3A_2$, d'où $A_2 = 2$. Si l'on fait $x = 0$, on a $2 = -A_1 + A_2 + N$, d'où $N = A_1$. Ces résultats simplifient beaucoup la résolution du système (6). Si une telle situation se présente en pratique, il faut en tirer parti.

REMARQUE 2. En principe, l'intégrale de toute fraction rationnelle peut être exprimée par des fonctions élémentaires. En pratique on ne peut intégrer entièrement (1) que si l'on connaît toutes les racines de $Q_n$ et leurs multiplicités. On a signalé au § 5 que cela n'était pas toujours facile. Aussi aurons-nous besoin de divers artifices pour simplifier l'intégration de la fraction rationnelle (1).

De ce point de vue la méthode d'Ostrogradski [1]) présente un intérêt indéniable.

## § 7. Intégration de fonctions irrationnelles

L'intégration de fonctions simples non rationnelles soulève de gros problèmes même si les intégrales de ces fonctions existent.

Nous allons traiter des cas où un changement de variable permet de ramener l'intégration d'une fonction irrationnelle à celle d'une fonction rationnelle.

Soit $R(x, y)$ une fonction rationnelle de $x$ et de $y$, c'est-à-dire une fonction que l'on obtient par des opérations arithmétiques sur $x$ et $y$. Par exemple

$$R(x, y) = \frac{axy + y^2}{cx + x^{10}y} \text{ est une fonction rationnelle,}$$

$$f(x, y) = \sqrt{x+y} + x^2 \text{ est irrationnelle.}$$

I. Calculer $\int R\left(x, \sqrt[m]{\frac{ax+b}{cx+d}}\right) dx$, où $a$, $b$, $c$ et $d$ sont des constantes, $m$, un entier naturel, $ad - bc \neq 0$, $R(x, y)$, une fonction rationnelle.

---

[1]) V. Ostrogradski, illustre mathématicien russe (1801-1861).

Une fonction de la forme $R\left(x, \sqrt[m]{\frac{ax+b}{cx+d}}\right)$ s'appelle *fonction homographique irrationnelle.*

Montrons qu'on peut ramener son intégration à celle d'une fonction rationnelle par le changement $t=\sqrt[m]{\frac{ax+b}{cx+d}}$. En effet, $t^m = \frac{ax+b}{cx+d}$, d'où $x=\frac{b-dt^m}{ct^m-a}$ est une fonction rationnelle de $t$. D'autre part,

$$dx=\frac{mt^{m-1}[ad-bc]}{(ct^m-a)^2}\,dt.$$

Donc

$$\int R\left(x, \sqrt[m]{\frac{ax+b}{cx+d}}\right)dx=$$
$$=\int R\left(\frac{b-dt^m}{ct^m-a},\ t\right)\frac{mt^{m-1}[ad-bc]}{(ct^m-a)^2}\,dt=\int R_1(t)\,dt,$$

où $R_1(t)$ est une fonction rationnelle de $t$ que nous savons intégrer.

Exemple 1. Calculer $\int \frac{1}{(x-1)^2}\sqrt[3]{\frac{x+1}{x-1}}\,dx$. Ici

$$R(x, y)=\frac{y}{(x-1)^2}.$$

En posant

$$\sqrt[3]{\frac{x+1}{x-1}}=t,$$

on obtient

$$x=\frac{t^3+1}{t^3-1},\qquad dx=\frac{-6t^2\,dt}{(t^3-1)^2},\qquad x-1=\frac{2}{t^3-1}.$$

Donc

$$\int \frac{1}{(x-1)^2}\sqrt[3]{\frac{x+1}{x-1}}\,dx=\int t\,\frac{-6t^2}{(t^3-1)^2}\cdot\frac{(t^3-1)^2}{4}\,dt=$$
$$=-\frac{3}{2}\int t^3\,dt=-\frac{3}{8}t^4+C=-\frac{3}{8}\left(\sqrt[3]{\frac{x+1}{x-1}}\right)^4+C.$$

Exemple 2.

$$\int\frac{dx}{\sqrt{x}+\sqrt[3]{x}}=\int\frac{dx}{(\sqrt[6]{x})^3+(\sqrt[6]{x})^2}=(\sqrt[6]{x}=t)=$$
$$=\int\frac{6t^5\,dt}{t^2+t^3}=6\int(t^2-t+1)\,dt-\ln|1+t|=$$
$$=2t^3-3t^2+6t-\ln|1+t|+C.$$

II. Calculer $\int R(x, \sqrt{ax^2+bx+c})\,dx$, où $a$, $b$ et $c$ sont des constantes.

Si le trinôme $ax^2+bx+c$ possède des racines réelles $x_1$ et $x_2$, alors $ax^2+bx+c=a(x-x_1)(x-x_2)$ et l'intégration de

$$R(x, \sqrt{ax^2+bx+c}) = R\left(x, (x-x_1)\sqrt{\frac{x-x_2}{x-x_1}a}\right) = R_1\left(x, \sqrt{\frac{x-x_2}{x-x_1}}\right)$$

se ramène au cas I.

On admettra donc que $ax^2+bx+c$ ne possède pas de racines réelles et $a>0$. On ramène l'intégration à celle d'une fonction rationnelle au moyen de la substitution d'Euler [1]:

$$t=\sqrt{ax^2+bx+c}+x\sqrt{a}.$$

D'où

$$ax^2+bx+c=t^2-2x\sqrt{a}\,t+ax^2,$$

c'est-à-dire que

$$x=\frac{t^2-c}{2t\sqrt{a}+b}$$

est une fonction rationnelle de $t$. Donc, il en est de même de

$$\sqrt{ax^2+bx+c}=t-x\sqrt{a}=t-\frac{t^2-c}{2t\sqrt{a}+b}\sqrt{a}$$

et

$$\int R(x, \sqrt{ax^2+bx+c})\,dx=\int R_1(t)\,dt.$$

Remarque. Si $a<0$ et $c>0$ ($ax^2+bx+c\geqslant 0$), on fait la substitution

$$\sqrt{ax^2+bx+c}=xt+\sqrt{c}.$$

Exemple 3. Calculer $\int \sqrt{a^2+x^2}\,dx$.

Le binôme ne possède pas de racines réelles. Donc on pose

$$t=\sqrt{x^2+a^2}+x,\quad x^2+a^2=t^2-2tx+x^2,\quad x=\frac{t^2-a^2}{2t}$$

et

$$\sqrt{x^2+a^2}=t-x=\frac{t^2+a^2}{2t}.$$

[1]) On peut utiliser cette substitution dans le cas de racines réelles pour $a>0$ sur l'intervalle où $ax^2+bx+c\geqslant 0$.

D'où

$$x\sqrt{x^2+a^2}=\frac{t^4-a^4}{4t^2}, \quad dx=\frac{t^2+a^2}{2t^2}\,dt.$$

Ainsi donc,

$$\int \sqrt{x^2+a^2}\,dx=\int \frac{t^2+a^2}{2t}\cdot\frac{t^2+a^2}{2t^2}\,dt=\frac{1}{4}\int\left[t+\frac{2a^2}{t}+\frac{a^4}{t^3}\right]dt=$$

$$=\frac{a^2}{2}\ln|t|+\frac{t^2}{8}-\frac{a^4}{8t^2}+C=\frac{a^2}{2}\ln|t|+\frac{t^4-a^4}{8t^2}+C=$$

$$=\frac{a^2}{2}\ln|x+\sqrt{x^2+a^2}|+\frac{x}{2}\sqrt{x^2+a^2}+C.$$

III. Intégration d'expressions de la forme $R(\cos x, \sin x)$.

On ramène l'intégration à celle d'une fonction rationnelle par la substitution $t=\operatorname{tg}(x/2)$ ($x\in\,]-\pi, \pi[$) dite *substitution universelle*. En effet,

$$\sin x=\frac{2\operatorname{tg}(x/2)}{1+\operatorname{tg}^2(x/2)}=\frac{2t}{1+t^2}, \quad \cos x=\frac{1-\operatorname{tg}^2(x/2)}{1+\operatorname{tg}^2(x/2)}=\frac{1-t^2}{1+t^2},$$

$$x=2\operatorname{Arc\,tg} t, \quad dx=\frac{2\,dt}{1+t^2},$$

donc

$$\int R(\cos x, \sin x)\,dx=\int R\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)\frac{2\,dt}{1+t^2}=\int R_1(t)\,dt.$$

Si la fonction $R(x, y)$ est paire ou impaire en $x$ ou en $y$, on peut se servir d'autres substitutions.

Soit

$$R(u, v)=\frac{P(u, v)}{Q(u, v)} \quad (u=\cos x, \ v=\sin x),$$

où $P$ et $Q$ sont des polynômes de $u$ et de $v$.

1) Si l'un des polynômes $P$ et $Q$ est pair en $v$ et l'autre impair en $v$, on fait la substitution $t=\cos x$.

2) Si l'un des polynômes $P$ et $Q$ est pair en $u$ et l'autre impair en $u$, on utilise la substitution $t=\sin x$.

3) Si $P$ et $Q$: a) ne changent pas quand on remplace $u$ et $v$ respectivement par $-u$ et $-v$, ou b) changent tous deux de signe, on se sert alors de la substitution $t=\operatorname{tg} x$ (ou $t=\operatorname{cotg} x$).

EXEMPLES:

1. $$\int\frac{dx}{\sin x}=\left(t=\operatorname{tg}\frac{x}{2}\right)=\int\frac{dt}{t}=\ln|t|+C=\ln\left|\operatorname{tg}\frac{x}{2}\right|+C.$$

2. $$\int\frac{\sin^3 x}{\cos^4 x}\,dx=-\int\frac{\sin^2 x\,d(\cos x)}{\cos^4 x}=-\int\frac{1-\cos^2 x}{\cos^4 x}\,d(\cos x)=$$

$$=(t=\cos x)=-\int\frac{1-t^2}{t^4}\,dt.$$

On a $R(u, v) = \frac{v^3}{u^1} = \frac{v^3}{u^1 v^0}$, c'est-à-dire que le numérateur est impair en $v$ et le dénominateur, pair en $v$. On a affaire au cas 1).

3. $$\int \frac{dx}{a^2 \cos^2 x + b^2 \sin^2 x} = \int \frac{dx}{(a^2 + b^2 \operatorname{tg}^2 x) \cos^2 x} =$$
$$= (t = \operatorname{tg} x) = \int \frac{dt}{a^2 + b^2 t^2}.$$

On a $P(u, v) = 1$ et $Q(u, v) = a^2 u^2 + b^2 v^2$. On se trouve dans le cas 3a), car $P$ et $Q$ ne changent pas quand on remplace $u$ et $v$ respectivement par $-u$ et $-v$.

CHAPITRE 6

# INTÉGRALE DÉFINIE

## § 1. Problèmes introduisant la notion d'intégrale définie. Définition de l'intégrale définie

a) Soit donnée sur un intervalle $[a, b]$ ($a$ et $b$ sont finis) une fonction $f(x)$ continue positive, dont le graphe est représenté sur la figure 75. Posons le problème suivant: définir la notion d'aire de la

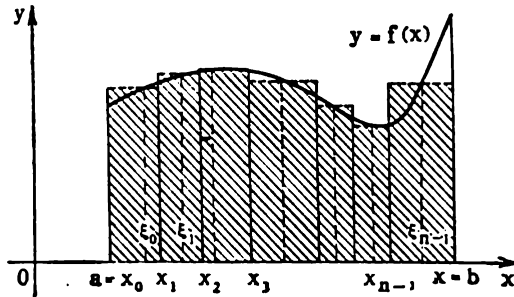



Fig. 75

figure limitée par la courbe $y = f(x)$, l'axe $Ox$ et les droites $x = a$ et $x = b$, et calculer cette aire. Ce problème se résout de la manière suivante. Divisons $[a, b]$ en $n$ parties par les points

$$a = x_0 < x_1 < \ldots < x_n = b, \tag{1}$$

choisissons dans chaque intervalle partiel

$$[x_j, x_{j+1}] \quad (j = 0, 1, \ldots, n-1) \tag{2}$$

un point arbitraire $\xi_j$, calculons les valeurs de $f$ aux points $\xi_j$ et formons la somme:

$$S_n = \sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j \quad (\Delta x_j = x_{j+1} - x_j) \tag{3}$$

que nous appellerons *somme de Riemann*. Cette somme est de toute évidence égale à celle de tous les rectangles hachurés (cf. fig. 75).

Faisons tendre tous les $\Delta x_j$ vers 0 de telle sorte que le plus grand $\Delta x_j$ tende aussi vers 0. Si la quantité $S_n$ tend vers une limite $S$ ne dépendant pas de la subdivision (1) et du choix des points $\xi_j$, il est naturel d'appeler la quantité $S$ *aire de la figure mixtiligne*. Donc,

$$S = \lim_{\max \Delta x_j \to 0} \sum_{j=0}^{n-1} f(\xi_j) \Delta x_j \qquad (4)$$

Nous avons ainsi défini l'aire de la figure mixtiligne. Il se pose aussitôt la question de savoir si une telle figure possède une aire, autrement dit de savoir si la somme de Riemann $S_n$ tend vers une limite finie pour $\Delta x_j \to 0$. On montrera dans la suite que la réponse est positive : toute figure mixtiligne correspondant à une fonction continue $f(x)$ possède bien une aire au sens défini plus haut, exprimée par un nombre $S$ dépendant de cette figure.

La deuxième question concerne l'adéquation d'une telle définition de l'aire. Disons que la pratique a complètement justifié cette définition et nous aurons d'ailleurs maintes occasions de nous en convaincre.

b) On demande la masse d'une tige non homogène portée par un axe $Ox$ et dont une extrémité se trouve au point d'abscisse $a$ et l'autre au point d'abscisse $b$ $(b > a)$.

Supposons que la densité de répartition de la masse le long de la tige est une fonction continue de $x$ : $\rho(x)$. Soit une subdivision de l'intervalle $[a, b]$ : $a = x_0 < x_1 < \ldots < x_n = b$. Choisissons un point $\xi_i$ quelconque dans chaque intervalle partiel $[x_i, x_{i+1}]$.

La fonction $\rho(x)$ variant peu sur l'intervalle $[x_i, x_{i+1}]$, on peut admettre que la masse de la portion $[x_i, x_{i+1}]$ de la tige est approximativement égale à $\rho(\xi_i) \Delta x_i$, où $\Delta x_i = x_{i+1} - x_i$.

La masse totale $m$ de la tige est approximativement égale à

$$\rho(\xi_0) \Delta x_0 + \rho(\xi_1) \Delta x_1 + \ldots + \rho(\xi_{n-1}) \Delta x_{n-1} = \sum_{i=0}^{n-1} \rho(\xi_i) \Delta x_i.$$

La masse exacte est de toute évidence la limite de cette expression lorsque le plus grand intervalle partiel tend vers zéro, soit

$$m = \lim_{\max \Delta x_i \to 0} \sum_{i=0}^{n-1} \rho(\xi_i) \Delta x_i. \qquad (4')$$

Les deux problèmes envisagés nous ont conduits à une même opération mathématique sur des fonctions d'origine différente définies sur un intervalle $[a, b]$. Nous rencontrerons beaucoup d'autres problèmes concrets dont la résolution se ramènera à une opération identique sur une fonction définie sur un intervalle fermé. Cette opération s'appelle *intégration d'une fonction sur un intervalle fermé*. Le nombre ainsi obtenu s'appelle *intégrale définie d'une fonction sur un intervalle fermé*.

DÉFINITION 1. *Soit donnée une fonction $f$ sur un intervalle* $[a, b]$. *Partageons* $[a, b]$ *en intervalles partiels avec les points arbitraires*

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

*On dira qu'on a effectué une subdivision $R$ de l'intervalle* $[a, b]$. *Prenons un point arbitraire* $\xi_j \in [x_j, x_{j+1}]$ *et formons la somme*

$$\sigma_R = \sigma_R(f) = \sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j \quad (\Delta x_j = x_{j+1} - x_j),$$

*appelée somme de Riemann de la fonction $f$ attachée à la subdivision $R$.*

*Désignons par*

$$\lambda_R = \max_{0 \leqslant j \leqslant n-1} \Delta x_j$$

*et appelons pas ou module de $R$ le plus grand des intervalles partiels* $[x_j, x_{j+1}]$.

*La limite (si elle existe) de la somme de Riemann $\sigma_R$ quand $\lambda_R \to 0$ s'appelle intégrale définie de la fonction $f$ sur l'intervalle* $[a, b]$ *et se note*

$$\lim_{\lambda_R \to 0} \sigma_R = \lim_{\max \Delta x_j \to 0} \sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j = \int_a^b f(x)\,dx \quad (a < b). \qquad (5)$$

Le nombre $a$ s'appelle *borne inférieure*, le nombre $b$, *borne supérieure* de l'intégrale définie.

La définition 1 est équivalente à la suivante.

DÉFINITION 1'. *On appelle intégrale définie d'une fonction $f$ sur un intervalle* $[a, b]$ *le nombre $I$ possédant la propriété suivante: pour tout $\varepsilon > 0$ on peut exhiber un nombre $\delta > 0$ tel que pour toute subdivision $R$ de* $[a, b]$ *de pas*

$$\lambda_R = \max_j \Delta x_j < \delta$$

*on ait*

$$|\sigma_R - I| = \left| \sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j - I \right| < \varepsilon$$

*quel que soit* $\xi_j \in [x_j, x_{j+1}]$.

La notion d'intégrale définie telle que nous venons de la définir a été introduite pour les fonctions continues par Cauchy et dans le cas général, c'est-à-dire pour les fonctions non nécessairement continues, par Riemann [1]). La limite (5) s'appelle *intégrale de Rie-*

[1]) Georg Friedrich Bernhard Riemann, éminent mathématicien allemand (1826-1866).

*mann* et la fonction pour laquelle cette limite existe *fonction intégrable-Riemann.*

Si une fonction $f$ est continue sur $[a, b]$, on verra que la limite (5) existe toujours pour elle.

On dit aussi qu'une fonction continue sur $[a, b]$ est *intégrable-Cauchy* sur $[a, b]$.

En a) nous avons défini (cf. fig. 75) l'aire d'une figure mixtiligne limitée par le graphe d'une fonction continue $y = f(x) \geqslant 0$, l'axe $Ox$ et les droites $x = a$ et $x = b$. Nous pouvons dire maintenant que l'aire de cette figure est égale à l'intégrale définie de $f$ sur $[a, b]$:

$$S = \int_a^b f(x)\, dx.$$

Nous pouvons dire encore que la masse de la tige étudiée en b) est égale à l'intégrale définie de la densité linéaire $\rho(x)$ sur $[a, b]$:

$$m = \int_a^b \rho(x)\, dx.$$

Ainsi, *l'intégrale définie d'une fonction $f$ sur un intervalle* $[a, b]$ *est par définition la limite de la somme de Riemann* (5) *lorsque le pas de la subdivision $R$ tend vers zéro.*

Dans cette définition qui désormais n'est plus liée à la recherche d'une aire, la fonction $f$ n'est pas nécessairement continue et positive sur $[a, b]$. Signalons que cette définition n'affirme pas l'existence de l'intégrale définie de toute fonction $f$ définie sur $[a, b]$, c'est-à-dire l'existence de la limite (5). Elle dit seulement que si cette limite existe pour une fonction donnée sur $[a, b]$, elle s'appelle intégrale de $f$ sur $[a, b]$.

A noter encore que lorsqu'on dit que la limite $I$ existe, on sous-entend qu'elle ne dépend pas du pas de la subdivision de $[a, b]$ et des points $\xi_j$.

Le calcul direct d'une intégrale définie à l'aide de la formule (5) soulève de grosses difficultés dans la mesure où les sommes de Riemann de fonctions parfois peu complexes prennent des proportions démesurées qui compliquent la détermination de ces limites. En tous les cas on n'a pas encore réussi à mettre au point des méthodes générales de calcul dans ce domaine. Il est intéressant de noter que c'est Archimède qui le premier a résolu des problèmes de cette nature. Il a en effet calculé l'aire d'un segment de parabole par des raisonnements qui rappellent vaguement la méthode actuelle des limites. Au fil des siècles de nombreux mathématiciens se sont penchés sur le calcul d'aires et de volumes de figures. Au XVII[e] siècle

encore la position de ces problèmes et les méthodes de leur résolution revêtaient un caractère particulier. Il faut attendre Newton et Leibniz [1]) pour voir enfin se dessiner une méthode générale de résolution. Ils ont montré que le calcul de l'intégrale définie d'une fonction pouvait se ramener à la recherche d'une primitive de cette fonction.

On a indiqué plus haut qu'une fonction continue sur $[a, b]$ était intégrable sur $[a, b]$. Ce fait sera justifié au § 7.

On démontrera de même qu'une fonction monotone sur $[a, b]$ est intégrable sur $[a, b]$. On rappelle qu'une fonction monotone peut

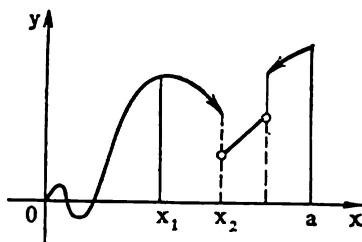



Fig. 76

posséder des discontinuités en nombre fini ou dénombrable (voir théorème 2 du § 4, chap. 3).

La figure 76 représente le graphe d'une fonction $y = f(x)$ définie sur un intervalle $[0, a]$. Cette fonction est continue sur $[0, x_1]$, décroissante sur $[x_1, x_2]$ et croissante sur $[x_2, a]$. Elle est donc intégrable sur chacun de ces intervalles et par suite sur l'intervalle $[0, a]$ tout entier en vertu de l'additivité de l'intégrale (voir § 2, théorème 3).

Donc, si un intervalle $[a, b]$ de définition d'une fonction $y = f(x)$ peut être partagé en un nombre fini d'intervalles partiels sur chacun desquels cette fonction est continue ou monotone, alors elle est intégrable sur $[a, b]$.

Newton et Leibniz ont prouvé un théorème liant deux notions fondamentales de l'analyse mathématique : l'intégrale et la dérivée. Ce théorème s'exprime par la relation (dite formule de Newton-Leibniz)

$$F(b) - F(a) = \int_a^b f(x)\, dx, \tag{6}$$

[1]) Isaac Newton, génial mathématicien et physicien anglais (1643-1727). Gottfried Wilhelm Leibniz, illustre mathématicien allemand (1646-1716).

où $f(x)$ est une fonction continue sur $[a, b]$, $F(x)$, une primitive de $f(x)$ sur $[a, b]$.

En résumé, *pour calculer l'intégrale définie d'une fonction continue sur un intervalle* $[a, b]$, *il faut connaître une primitive* $F(x)$ *de* $f$ *et calculer la différence* $F(b) - F(a)$ *des valeurs prises par* $F(x)$ *aux extrémités de l'intervalle* $[a, b]$.

On déduit sans peine la formule (6) si l'on sait qu'une fonction $f(x)$ continue sur $[a, b]$ est intégrable sur $[a, b]$ et qu'elle possède une primitive $F(x)$.

Soit $R$ une subdivision quelconque

$$a = x_0 < x_1 < \ldots < x_n = b$$

de l'intervalle $[a, b]$. Alors (voir justifications plus bas)

$$F(b) - F(a) = F(x_n) - F(x_0) = F(x_n) - F(x_{n-1}) + F(x_{n-1}) - \ldots$$
$$\ldots - F(x_1) + F(x_1) - F(x_0) = \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k)] =$$
$$= \sum_{=0}^{n-1} F'(\xi_k)(x_{k+1} - x_k) = \sum_{k=0}^{n-1} f(\xi_k)\,\Delta x_k \xrightarrow[\lambda_R \to 0]{} \int_a^b f(x)\,dx, \quad (7)$$

d'où la formule (6).

Dans la quatrième égalité (7) on s'est servi du théorème des accroissements finis

$$F(x_{k+1}) - F(x_k) = F'(\xi_k)(x_{k+1} - x_k),$$

où $\xi_k \in ]x_k, x_{k+1}[$. La dernière relation résulte du fait que la fonction $f$ étant continue sur $[a, b]$ est intégrable sur $[a, b]$ et par suite l'une quelconque de ses sommes de Riemann, et en particulier celle obtenue en appliquant le théorème des accroissements finis, tend vers l'intégrale définie de $f$ sur $[a, b]$ lorsque $\lambda_R \to 0$.

On a le

THÉORÈME. *Une fonction non bornée sur un intervalle* $[a, b]$ *n'est pas intégrable sur cet intervalle.*

Donc, *pour qu'une fonction* $f$ *soit intégrable sur un intervalle* $[a, b]$ *il est nécessaire qu'elle soit bornée sur cet intervalle.*

Cependant, cette condition n'est pas suffisante.

EXEMPLE. La fonction

$$\psi(x) = \begin{cases} 1 \text{ si } x \text{ est rationnel,} \\ -1 \text{ si } x \text{ est irrationnel,} \end{cases}$$

est bornée: $|\psi(x)| = 1$, mais pas intégrable sur tout intervalle $[a, b]$.

En effet, si les points $\xi_j$ figurant dans la somme de Riemann de $\psi(x)$ sont rationnels, alors

$$\sigma_R = \sum_{j=0}^{n-1} \psi(\xi_j)\,\Delta x_j = \sum_{j=0}^{n-1} 1\cdot\Delta x_j = b-a.$$

Si les points $\xi_j$ sont irrationnels, alors

$$\sigma_R = \sum_{j=0}^{n-1} (-1)\,\Delta x_j = -(b-a).$$

La somme $\sigma_R$ n'étant pas la même dans les deux cas, la fonction n'est pas intégrable sur $[a, b]$.

DÉMONSTRATION DU THÉORÈME. Soit

$$\sigma_R = \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1}-x_i)$$

une somme de Riemann de la fonction $f$ attachée à la subdivision $R$: $a = x_0 < x_1 < \ldots < x_n = b$. Si l'on admet que la fonction $f$ n'est pas bornée sur $[a, b]$, elle le sera nécessairement sur un intervalle partiel, pour fixer les idées, $[x_{i_0}, x_{i_0+1}]$. Figeons $\xi_i \in [x_i, x_{i+1}]$ pour tous les $i \neq i_0$ et supposons provisoirement que $\xi_{i_0}$ est variable. Le terme $f(\xi_{i_0})(x_{i_0+1}-x_{i_0})$ n'est pas borné sur $[x_{i_0}, x_{i_0+1}]$, quant à la somme des autres termes, elle est égale à un nombre bien défini. Donc, $|\sigma_R|$ peut être rendue arbitrairement grande par un choix convenable du point $\xi_{i_0}$ et la fonction $f$ ne sera pas intégrable sur $[a, b]$. En effet, si une fonction est intégrable sur $[a, b]$, ses sommes de Riemann sont bornées quel que soit $\xi_i$.

Dans la suite, on introduira la notion d'intégrale impropre et l'on verra que certaines fonctions non bornées sur un intervalle fermé borné admettent des intégrales impropres.

## § 2. Propriétés des intégrales définies

On étudie dans ce paragraphe les propriétés des fonctions intégrables. On a signalé plus haut que les fonctions continues et monotones sur un intervalle $[a, b]$ étaient intégrables sur cet intervalle. Ce fait sera prouvé au § 7.

THÉORÈME 1 *Si $M$ est une constante, alors*

$$\int_a^b M\,dx = M(b-a). \tag{1}$$

En effet, quelle que soit la subdivision $R$ de $[a, b]$, la somme de Riemann de la fonction $f(x) = M$ est égale à

$$\sigma_R = \sum_{j=0}^{n-1} M\,\Delta x_j = M \sum_{j=0}^{n-1} \Delta x_j = M(b-a).$$

D'où

$$\lim_{\lambda_R \to 0} \sigma_R = M(b-a).$$

THÉORÈME 2. *Pour la fonction*

$$\psi_c(x) = \begin{cases} 0, & x \in [a, b],\ x \neq c, \\ A, & x = c, \end{cases}$$

*on a*

$$\int_a^b \psi_c(x)\,dx = 0.$$

En effet, considérons une subdivision $R$ de $[a, b]$:

$$a = x_0 < x_1 < \ldots < x_n = b.$$

L'un des intervalles semi-ouverts, disons $[x_m, x_{m+1}[$, contient le point $c$. Donc

$$\sigma_R(\psi_c) = \sum_{k=0}^{n-1} \psi_c(\xi_k)\,\Delta x_k = \psi_c(\xi_{m-1})\,\Delta x_{m-1} + \psi_c(\xi_m)\,\Delta x_m$$

(les autres termes sont manifestement nuls). Comme $|\psi_c(x)| \leqslant |A|$, alors

$$|\sigma_R(\psi_c)| \leqslant |A|\,(\Delta x_{m-1} + \Delta x_m) \to 0$$

pour $\lambda_R \to 0$, c.q.f.d.

THÉORÈME 3. *Si une fonction $f$ est intégrable sur chacun des intervalles $[a, c]$, $[c, b]$ $(a < c < b)$, elle le sera sur $[a, b]$ et*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \tag{2}$$

(*additivité de l'intégrale définie*).

DÉMONSTRATION. Soit une subdivision $R$ de $[a, b]$:

$$R:\ a = x_0 < x_1 < \ldots < x_n = b.$$

On admettra que $c$ est confondu avec un point de $R$: $x_m = c$. Alors $R$ induit sur $[a, c]$ et $[c, b]$ les subdivisions $R_1$ et $R_2$:

$$R_1:\ a = x_0 < x_1 < \ldots < x_m = c,$$

$$R_2:\ c = x_m < x_{m+1} < \ldots < x_n = b,$$

et

$$\sigma_R=\sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j=\sum_{j=0}^{m-1} f(\xi_j)\,\Delta x_j+\sum_{j=m}^{n-1} f(\xi_j)\,\Delta x_j=\sigma_{R_1}+\sigma_{R_2}.$$

Supposons que

$$\lambda_R=\max_{0\leqslant j\leqslant n-1} |\Delta x_j| \to 0.$$

A fortiori $\lambda_{R_1}\to 0$ et $\lambda_{R_2}\to 0$. Et par suite

$$\lim_{\lambda_R\to 0}\sigma_R=\lim_{\lambda_{R_1}\to 0}\sigma_{R_1}+\lim_{\lambda_{R_2}\to 0}\sigma_{R_2}=\int_a^c f(x)\,dx+\int_c^b f(x)\,dx.$$

Cette relation a été établie pour des $R$ contenant le point $c$. Elle est alors valable pour toute $R$ (voir lemme plus bas). Donc, l'intégrale $\int_a^b f(x)\,dx$ existe et l'on a (2).

Par définition,

$$\int_a^a f(x)\,dx=0, \tag{3}$$

$$\int_a^b f(x)\,dx=-\int_b^a f(x)\,dx, \tag{4}$$

où $f$ est intégrable sur $[a, b]$.

Il est immédiat de voir que, eu égard à (3) et à (4), la relation (2) est valable quels que soient $a$, $b$ et $c$, pourvu que $f$ soit intégrable sur le plus grand des intervalles $[a, b]$, $[a, c]$, $[c, b]$.

Si par exemple $c<a<b$, le théorème 3 nous dit que

$$\int_c^b f\,dx=\int_c^a f\,dx+\int_a^b f(x)\,dx$$

ou

$$\int_a^b f\,dx=\int_c^b f\,dx-\int_c^a f\,dx=\int_a^c f\,dx+\int_c^b f\,dx,$$

et l'on obtient (2).

THÉORÈME 4. *Si des fonctions $f_1$ et $f_2$ sont intégrables sur $[a, b]$ et $A$ et $B$ sont des nombres arbitraires, alors*

$$\int_a^b (Af_1+Bf_2)\,dx=A\int_a^b f_1\,dx+B\int_a^b f_2\,dx. \tag{5}$$

*En particulier, pour $B = 0$, on obtient*

$$\int_a^b Af_1\,dx = A\int_a^b f_1\,dx, \tag{6}$$

*relation qui exprime qu'on peut sortir un facteur constant du signe d'intégration.*

*Pour $A = 1$ et $B = \pm 1$, on obtient*

$$\int_a^b (f_1 \pm f_2)\,dx = \int_a^b f_1\,dx \pm \int_a^b f_2\,dx. \tag{7}$$

DÉMONSTRATION. Pour une subdivision $R$ arbitraire, on a

$$\sum_{j=0}^{n-1} [Af_1(\xi_j) + Bf_2(\xi_j)]\,\Delta x_j = A\sum_{j=0}^{n-1} f_1(\xi_j)\,\Delta x_j + B\sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j.$$

De là, en passant à la limite pour $\lambda_R \to 0$, on déduit l'égalité (5) qui, de toute évidence, est valable pour $b \leqslant a$ aussi.

THÉORÈME 5. *Une fonction $f$ étant intégrable sur $[a, b]$, si l'on pose $f_1(x) = f(x) + \psi_c(x)$, où $c \in [a, b]$ et*

$$\psi_c(x) = \begin{cases} 0 \text{ pour } x \in [a, b],\ x \neq c, \\ A \text{ pour } x = c \ (A \textit{ est un nombre quelconque}), \end{cases}$$

*alors*

$$\int_a^b f_1(x)\,dx = \int_a^b f(x)\,dx.$$

DÉMONSTRATION. Le théorème 2 nous dit que

$$\int_a^b \psi_c(x)\,dx = 0.$$

Donc, en vertu du théorème 4,

$$\int_a^b f_1(x)\,dx = \int_a^b f(x)\,dx + \int_a^b \psi_c(x)\,dx = \int_a^b f(x)\,dx,$$

c.q.f.d.

REMARQUE 1. Le théorème 5 nous dit que l'intégrabilité d'une fonction $f$ ne dépend pas des valeurs qu'elle prend en un point déterminé.

Ainsi, la fonction $\psi(x) = (\sin x)/x$ est définie sur l'intervalle $]0, 1]$. Si l'on pose $\psi(0) = 1$, elle sera continue et par suite intégrable sur l'intervalle $[0, 1]$. Cependant, que l'on pose $\psi(0) = 1$ ou

$\psi(0) = A$, où $A$ est un nombre quelconque, la fonction $\psi(x)$ restera intégrable et son intégrale $\int_0^1 \psi(x)\,dx$ prendra la même valeur.

THÉORÈME 6. *Si des fonctions $f$ et $\varphi$ intégrables sur un intervalle $[a, b]$ sont telles que*

$$f(x) \leqslant \varphi(x),$$

*alors*

$$\int_a^b f(x)\,dx \leqslant \int_a^b \varphi(x)\,dx \quad (a \leqslant b). \tag{8}$$

DÉMONSTRATION. Pour toute subdivision $R$, on a

$$\sum_{j=0}^{n-1} f(\xi_j)\,\Delta x_j \leqslant \sum_{j=0}^{n-1} \varphi(\xi_j)\,\Delta x_j,$$

car $\Delta x_j > 0$. En passant à la limite pour $\lambda_R \to 0$, on obtient (8).

THÉORÈME 7. *On a*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx \quad (a \leqslant b) \tag{9}$$

*ou si $b < a$*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \left|\int_a^b |f(x)|\,dx\right|, \tag{9'}$$

*pourvu que $f$ et $|f|$ soient intégrables sur $[a, b]$.*

DÉMONSTRATION. Il est évident que

$$-|f(x)| \leqslant f(x) \leqslant |f(x)|, \quad \forall x \in [a, b].$$

En vertu du théorème 6

$$\int_a^b (-|f(x)|)\,dx \leqslant \int_a^b f\,dx \leqslant \int_a^b |f|\,dx \quad (a < b)$$

ou

$$-\int_a^b |f|\,dx \leqslant \int_a^b f\,dx \leqslant \int_a^b |f|\,dx,$$

ou encore

$$\left|\int_a^b f\,dx\right| \leqslant \int_a^b |f|\,dx \quad (a<b),$$

c.q.f.d.

Si $a<b$, les seconds membres de (9) et (9′) sont égaux.
Si $b<a$, alors en vertu de (4), on a

$$\left|\int_a^b f\,dx\right| = \left|\int_b^a f\,dx\right| \leqslant \int_b^a |f|\,dx = \left|\int_a^b |f|\,dx\right|,$$

c'est-à-dire (9′).

Enfin, le cas $a=b$ se ramène à la relation évidente $0 \leqslant 0$. Ce qui prouve (9).

REMARQUE 2. L'intégrabilité de $f$ sur $[a, b]$ entraîne celle de $|f|$ sur $[a, b]$ (voir plus bas § 7, remarque 2). Ceci est toujours évident pour des fonctions concrètes. Si par exemple $f$ est une fonction continue par morceaux sur $[a, b]$ (on démontrera qu'elle est intégrable), alors $|f|$ l'est aussi.

La réciproque n'est pas généralement vraie. L'intégrabilité de $|f(x)|$ n'implique pas nécessairement celle de $f(x)$ sur $[a, b]$.

Ainsi, la fonction $\psi(x)$ de l'exemple du § 1 :

$$\psi(x) = \begin{cases} 1 & \text{pour } x \text{ rationnel,} \\ -1 & \text{pour } x \text{ irrationnel,} \end{cases}$$

n'est pas intégrable sur $[a, b]$, cependant que $|\psi(x)| = 1$ l'est sur $[a, b]$.

THÉORÈME 8. *Etant donnée une fonction $f$ intégrable et positive sur $[a, b]$, s'il existe un point $c \in [a, b]$ en lequel $f$ est continue et tel que $f(c) > 0$, alors*

$$\int_a^b f(x)\,dx > 0. \tag{10}$$

DÉMONSTRATION. On admettra que $c \in \,]a, b[$. Comme $f$ est continue en $c$ et $f(c) > 0$, il existe un intervalle $[c-\delta, c+\delta]$ tel que (voir chap. 3, § 3, théorème 4)

$$f(x) > \frac{f(c)}{2} = \eta > 0, \quad \forall x \in [c-\delta, c+\delta].$$

Alors

$$\int_a^b f(x)\,dx = \int_a^{c-\delta} f(x)\,dx + \int_{c-\delta}^{c+\delta} f(x)\,dx + \int_{c+\delta}^b f(x)\,dx > 0,$$

car

$$\int_a^{c-\delta} f(x)\,dx \geqslant 0, \qquad \int_{c+\delta}^{b} f(x)\,dx \geqslant 0,$$

$$\int_{c-\delta}^{c+\delta} f(x)\,dx \geqslant \int_{c-\delta}^{c+\delta} \eta\,dx = 2\delta\eta > 0.$$

Si $c = a$ ou $c = b$, on considérera respectivement les intervalles $[a, a+\delta]$ et $[b-\delta, b]$ au lieu de $[c-\delta, c+\delta]$.

LEMME. *Soit $R_*$ une subdivision de $[a, b]$ dont un point est confondu avec c. Si une fonction f est bornée sur $[a, b]$ et ses sommes de Riemann attachées aux subdivisions de la forme $R_*$ sont telles que*

$$\lim_{\lambda_{R_*}\to 0} \sigma_{R_*} = I,$$

*alors f est intégrable sur $[a, b]$ et*

$$\int_a^b f(x)\,dx = I = \lim_{\lambda_R\to 0} \sigma_R.$$

DÉMONSTRATION. Soit $R$ une subdivision quelconque de $[a, b]$:

$$R:\ a = x_0 < x_1 < \ldots < x_m < x_{m+1} < \ldots < x_n = b,$$

telle que $x_m < c < x_{m+1}$.

En ajoutant le point $c$ à $R$ on obtient une subdivision $R_*$. Si $\lambda_R \to 0$, il en est de même de $\lambda_{R_*}$.

Si de $\sigma_R$ on élimine le terme $f(\xi_m)(x_{m+1} - x_m)$ et que l'on ajoute $f(\xi'_m) \times (c - x_m) + f(\xi''_m)(x_{m+1} - c)$, on obtient la somme de Riemann $\sigma_{R_*}$. Ceci étant

$$\sigma_R = \sigma_{R_*} + \mu,$$

où

$$\mu = f(\xi_m)(x_{m+1} - x_m) - f(\xi'_m)(c - x_m) - f(\xi''_m)(x_{m+1} - c),$$
$$x_m \leqslant \xi'_m \leqslant c, \qquad c \leqslant \xi''_m \leqslant x_{m+1}.$$

Il est évident que

$$|\mu| \leqslant M(x_{m+1} - x_m) + M(c - x_m) + M(x_{m+1} - c) = 2M(x_{m+1} - x_m) \xrightarrow[\lambda_R\to 0]{} 0.$$

Donc

$$\lim_{\lambda_R\to 0} \sigma_R = \lim_{\lambda_{R_*}\to 0} \sigma_{R_*} + \lim_{x_{m+1}-x_m\to 0} \mu = I + 0 = I.$$

## § 3. Intégrale fonction de sa borne supérieure

On remarquera que

$$\int_a^b f(x)\,dx = \int_a^b f(u)\,du,$$

c'est-à-dire que l'intégration peut s'effectuer aussi bien par rapport à $x$ qu'à $u$ sur $[a, b]$. En effet, dans les deux cas toute somme de Riemann de $f$ s'écrit

$$\sigma_R = \sum_{j=0}^{n-1} f(\xi_j)\, \Delta x_j.$$

Soit donnée une fonction $f$ intégrable sur un intervalle $[a, b]$. La fonction $f$ est également intégrable sur $[a, x]$ quel que soit $x \in \in [a, b]$.

Cette proposition est à démontrer, mais nous ne le ferons pas. En règle générale, elle est évidente. Par exemple, une fonction (monotone) continue sur un intervalle $[a, b]$ l'est sur $[a, x]$ et par suite est intégrable sur $[a, x]$.

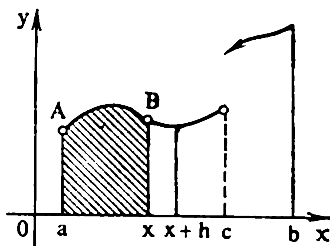


Fig. 77

Soit $x \in [a, b]$. Etudions l'intégrale définie de $f$ sur l'intervalle $[a, x]$. Ce sera une fonction de $x$ que nous désignerons par $F(x)$:

$$F(x) = \int_a^x f(u)\, du. \qquad (1)$$

La variable d'intégration est $u$ pour éviter toute confusion avec la borne supérieure $x$.

La figure 77 représente le graphe d'une fonction $f$ bornée, continue par morceaux, discontinue en $c$. Le nombre $F(x)$ est égal à l'aire de la figure $ABxa$. Il varie avec $x$.

THÉORÈME 1. *Si une fonction $f$ est intégrable sur un intervalle $[a, b]$, la fonction $F$ définie par la formule* (1) *est continue en tout point $x \in [a, b]$.*

DÉMONSTRATION. Prenons un point $x$ quelconque et donnons-lui un accroissement $h$ (fig. 77). On a

$$|F(x+h)-F(x)| = \left| \int_a^{x+h} f(u)\, du - \int_a^x f(u)\, du \right| = \left| \int_x^{x+h} f(u)\, du \right| \leqslant M\, |h|,$$

où $M \geqslant |f(u)|$, $\forall\, u \in [a, b]$.

L'inégalité obtenue

$$|F(x+h) - F(x)| \leqslant M\, |h|$$

implique que

$$\lim_{h \to 0} [F(x+h) - F(x)] = 0,$$

c'est-à-dire que $F$ est continue en $x$.

Signalons que $F(x)$ est continue en tout point $x$ qu'il soit point de continuité ou de discontinuité de $f$.

THEOREME 2. *Si une fonction $f$ intégrable sur $[a, b]$ est continue en un point $x \in [a, b]$, alors la fonction $F$ est dérivable en $x$* (voir (1)):

$$F'(x) = f(x). \tag{2}$$

DEMONSTRATION. Soit $x$ un point de continuité de $f$. On a

$$\frac{F(x+h)-F(x)}{h}=\frac{1}{h}\left[\int_a^{x+h} f(u)\,du-\int_a^x f(u)\,du\right]=\frac{1}{h}\int_x^{x+h} f(u)\,du=$$

$$=\frac{1}{h}\int_x^{x+h}\{f(x)+[f(u)-f(x)]\}\,du=$$

$$=\frac{1}{h}f(x)\,h+\frac{1}{h}\int_x^{x+h}[f(u)-f(x)]\,du=$$

$$=f(x)+\frac{1}{h}\int_x^{x+h}[f(u)-f(x)]\,du. \tag{3}$$

Pour obtenir (3) on s'est servi des propriétés de l'intégrale définie. Dans la quatrième égalité on a utilisé le fait que $f(x)$ ne dépendant pas de $u$ est considérée comme un facteur constant dans l'intégration par rapport à $u$ (voir théorème 1 du § 2). Prouvons que

$$\frac{1}{h}\int_x^{x+h}[f(u)-f(x)]\,du \xrightarrow[h\to 0]{} 0. \tag{4}$$

La fonction $f$ est continue en $x$, donc pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que

$$|f(u)-f(x)| < \varepsilon, \quad \forall u \in [x, x+h],$$

pourvu que $|h| < \delta$. Donc

$$\left|\frac{1}{h}\int_x^{x+h}[f(u)-f(x)]\,du\right|\leqslant\left|\frac{1}{h}\int_x^{x+h}|f(u)-f(x)|\,du\right|\leqslant$$

$$\leqslant\left|\frac{1}{h}\int_x^{x+h}\varepsilon\,du\right|=\left|\frac{1}{h}\varepsilon h\right|=\varepsilon,$$

ce qui prouve la propriété (4).

En passant maintenant à la limite pour $h \to 0$ dans (3), on établit l'existence de la dérivée $F'(x)$:

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = f(x).$$

Ceci prouve le théorème 2.

Signalons que dans le théorème 2 la fonction $f$ peut présenter des points de discontinuité sur $[a, b]$, mais au point $x$, où l'on affirme l'existence de la dérivée de $F$, la fonction $f$ doit nécessairement être continue, sinon le théorème 2 n'aurait pas de sens.

Le théorème 2 affirme en particulier que si une fonction $f(x)$ est continue sur un intervalle $[a, b]$, alors $F(x)$ possède sur cet intervalle une dérivée égale à $f(x)$.

Donc, *si une fonction $f$ est continue sur un intervalle* $[a, b]$, *elle possède une primitive sur cet intervalle. Pour primitive on peut prendre l'intégrale* (1).

Il s'ensuit que *l'intégrale indéfinie d'une fonction $f$ continue sur* $[a, b]$ *est égale à*

$$\int f(x)\,dx = \int_a^x f(u)\,du + C, \quad x \in [a, b],$$

*où $C$ est une constante.*

## § 4. Formule de Newton-Leibniz

Cette formule s'écrit

$$\int_a^b f(u)\,du = \Phi(b) - \Phi(a) = \Phi(x)\Big|_{x=a}^{x=b}, \tag{1}$$

$f(u)$ étant une fonction continue sur $[a, b]$, $\Phi(u)$, une primitive de $f$ sur $[a, b]$.

La formule de Newton-Leibniz a déjà été prouvée au § 1. On avait alors admis qu'une fonction continue sur $[a, b]$ était intégrable et possédait une primitive sur $[a, b]$.

On sait maintenant (voir § 3) que l'intégrabilité d'une fonction continue sur $[a, b]$ implique l'existence d'une primitive de cette fonction sur $[a, b]$.

Voici une autre démonstration de la formule de Newton-Leibniz. Revenons à la fonction

$$F(x) = \int_a^x f(u)\,du. \tag{2}$$

On remarquera que

$$F(a)=\int_a^a f(u)\,du=0 \quad \text{et} \quad F(b)=\int_a^b f(u)\,du. \tag{3}$$

On sait par ailleurs que $F(x)$ est une primitive de $f$ sur $[a, b]$. Donc,

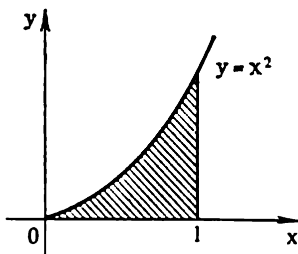



Fig. 78

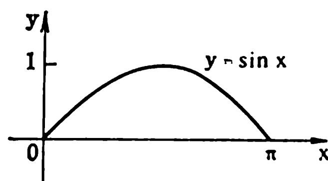



Fig. 79

si $\Phi(x)$ est une autre primitive de $f$, il existe une constante $C$ telle que

$$\Phi(x) = F(x) + C, \quad \forall x \in [a, b]. \tag{4}$$

De (2), (3) et (4) il s'ensuit

$$\Phi(b)-\Phi(a)=F(b)-F(a)=\int_a^b f(u)\,du,$$

ce qui prouve la formule (1).

Exemple 1.

$$\int_0^1 x^2\,dx=\frac{x^3}{3}\Big|_{x=0}^{x=1}=\frac{1}{3}.$$

Ceci montre que l'aire de la figure hachurée (fig. 78) est égale à 1/3.

Exemple 2.

$$\int_0^\pi \sin x\,dx=-\cos x\,|_0^\pi=1+1=2.$$

Donc, l'aire de la figure (fig. 79) limitée par l'arc de sinusoïde $y = \sin x$ et l'axe $Ox$ est égale à 2.

Exemple 3. La fonction

$$\varphi(x)=\operatorname{sgn} x=\begin{cases}-1, & -1\leqslant x<0,\\ 0, & x=0,\\ 1, & 0<x\leqslant 1,\end{cases}$$

est continue sur l'intervalle $[-1, 1]$ sauf en $x = 0$. L'intervalle $[-1, 1]$ peut être partagé en deux intervalles $[-1, 0]$ et $[0, 1]$ sur lesquels $\varphi$ est monotone, donc intégrable. On a la formule

$$F(x) = \int_{-1}^{x} \operatorname{sgn} u \, du = -1 + |x| \quad (-1 \leqslant x \leqslant 1). \tag{5}$$

En effet, la fonction $\varphi(x)$ étant continue sur $[-1, 0[$, sa primitive est égale à $-x$. En appliquant la formule de Newton-Leibniz, on obtient

$$\int_{-1}^{x} \operatorname{sgn} u \, du = \int_{-1}^{x} (-1) \, du = -u \Big|_{-1}^{x} = -1 - x \quad (-1 \leqslant x < 0) \tag{6}$$

Le théorème 1 du § 3 nous dit que la fonction $F(x)$ est continue, notamment en $x = 0$. Donc

$$F(0) = \lim_{x \to 0} (-1 - x) = -1. \tag{7}$$

Pour $x > 0$

$$F(x) = \int_{-1}^{x} \operatorname{sgn} u \, du = \int_{-1}^{0} \operatorname{sgn} u \, du + \int_{0}^{x} 1 \cdot du = -1 + u \Big|_{0}^{x} = -1 + x. \tag{8}$$

De (6), (7) et (8) on déduit (5).

On obtient une formule plus élégante en intégrant à partir de de $x = 0$:

$$\int_{0}^{x} \operatorname{sgn} u \, du = |x|. \tag{9}$$

L'intégrant de (9) est une fonction bornée, discontinue en $x = 0$. L'intégrale $F(x) = |x|$ comme fonction de sa borne supérieure est continue même en $x = 0$, ce qui est en accord avec le théorème 1 du § 3. Cependant, la dérivée $F'(0)$ n'existe pas, ce qui ne contredit pas le théorème 2 du § 3 qui affirme que $F'(x)$ existe si seulement $f$ est continue en $x$.

THÉORÈME 1 (DE CHANGEMENT DE LA VARIABLE). *On a*

$$\int_{a}^{b} f(x) \, dx = \int_{c}^{d} f[\varphi(t)] \, \varphi'(t) \, dt, \tag{10}$$

*où la fonction* $\varphi(t)$ *est continûment dérivable sur* $[c, d]$, $a = \varphi(c)$, $b = \varphi(d)$ *et* $f(x)$ *continue sur* $[A, B]$, *image de l'intervalle* $[c, d]$ *par* $\varphi$.

DÉMONSTRATION. Soient $F(x)$ et $\Phi(t)$ des primitives respectives de $f(x)$ et de $f[\varphi(t)] \, \varphi'(t)$. On a alors (voir chap. 5, § 2, (1) et plus

bas) $\Phi(t) = F[\varphi(t)] + C$, $c \leqslant t \leqslant d$, où $C$ est une constante. Donc,

$$F(b) - F(a) = F[\varphi(d)] - F[\varphi(c)] = \Phi(d) - \Phi(c). \qquad (11)$$

En vertu de la formule de Newton-Leibniz, le premier membre de (11) est égal au premier membre de (10) et le second membre de (11), au second membre de (10), ce qui prouve la formule (10).

EXEMPLE 4.

$$\int_0^a \sqrt{a^2-x^2}\,dx = (x = a\sin t) = \int_0^{\pi/2} \sqrt{a^2-a^2\sin^2 t}\,a\cos t\,dt =$$

$$= a^2\int_0^{\pi/2} \cos^2 t\,dt = a^2\int_0^{\pi/2} \frac{1+\cos 2t}{2}\,dt = \frac{a^2}{2}\left[t+\frac{\sin 2t}{2}\right]_{t=0}^{t=\frac{\pi}{2}} = \frac{\pi a^2}{4}.$$

REMARQUE. On peut prendre la borne supérieure d'intégration égale à $\frac{5}{2}\pi$.

EXEMPLE 5.

$$\int_0^{4\pi} \sin^3 t\,dt = -\int_0^{4\pi} (1-\cos^2 t)\,d(\cos t) = (x = \cos t) =$$

$$= -\int_1^1 (1-x^2)\,dx = 0,$$

car les bornes supérieure et inférieure sont égales.

EXEMPLE 6. Si $f$ est une fonction paire ($f(-u) = f(u)$), alors

$$\int_{-a}^{a} f(u)\,du = 2\int_0^a f(u)\,du,$$

car

$$\int_{-a}^{0} f(u)\,du = (u = -x) = -\int_a^0 f(-x)\,dx = \int_0^a f(-x)\,dx =$$

$$= \int_0^a f(x)\,dx = \int_0^a f(u)\,du.$$

EXEMPLE 7. Si $f$ est une fonction impaire ($f(-u) = -f(u)$), alors

$$\int_{-a}^{a} f(x)\,dx = 0.$$

EXEMPLE 8. Si $f$ est une fonction $2\pi$-périodique ($f(x+2\pi) = f(x)$), alors

$$\int_{\alpha}^{\alpha+2\pi} f(x)\,dx = \int_{0}^{2\pi} f(x)\,dx.$$

En effet,

$$\int_{2\pi}^{2\pi+\alpha} f(x)\,dx = (x = t + 2\pi) = \int_{0}^{\alpha} f(t+2\pi)\,dt = \int_{0}^{\alpha} f(t)\,dt = -\int_{\alpha}^{0} f(t)\,dt$$

et par suite

$$\int_{\alpha}^{2\pi+\alpha} f(x)\,dx = \int_{\alpha}^{0} f(x)\,dx + \int_{0}^{2\pi} f(x)\,dx + \int_{2\pi}^{2\pi+\alpha} f(x)\,dx = \int_{0}^{2\pi} f(x)\,dx.$$

EXEMPLE 9.

$$\int_{0}^{3\pi} \sin^3 t\,dt = (x = \cos t) = -\int_{1}^{-1} (1-x^2)\,dx = \int_{-1}^{1} (1-x^2)\,dx =$$

$$= 2\int_{0}^{1} (1-x^2)\,dx = 2\left(x - \frac{x^3}{3}\right)\Big|_0^1 = \frac{4}{3}.$$

EXEMPLE 10. Calculons l'intégrale de l'exemple 5 en se servant des résultats des exemples 8 et 7:

$$\int_{0}^{4\pi} \sin^3 t\,dt = \int_{0}^{2\pi} \sin^3 t\,dt + \int_{2\pi}^{4\pi} \sin^3 t\,dt = 2\int_{-\pi}^{\pi} \sin^3 t\,dt = 0,$$

puisque la fonction $\sin^3 t$ est impaire.

THÉORÈME 2. *On a la formule d'intégration par parties pour une intégrale définie*:

$$\int_{a}^{b} u'(x)\,v(x)\,dx = u(x)\,v(x)\big|_a^b - \int_{a}^{b} u(x)\,v'(x)\,dx, \qquad (12)$$

*où $u$ et $v$ sont des fonctions continûment dérivables sur* $[a, b]$.

DÉMONSTRATION. Le produit $u(x)\,v(x)$ possède une dérivée continue sur $[a, b]$:

$$(u(x)\,v(x))' = u(x)\,v'(x) + u'(x)\,v(x).$$

Donc, en vertu du théorème de Newton-Leibniz

$$u(x)\,v(x)\big|_a^b=\int_a^b [u(x)\,v'(x)+u'(x)\,v(x)]\,dx=$$

$$=\int_a^b u(x)\,v'(x)\,dx+\int_a^b u'(x)\,v(x)\,dx,$$

d'où l'on déduit (12).

EXEMPLE 11.

$$\int_0^1 \ln(1+x)\,dx=(u=\ln(1+x),\ dv=dx)=$$

$$=x\ln(1+x)\big|_0^1-\int_0^1\frac{x\,dx}{1+x}=\ln 2-\int_0^1 1\,dx+\int_0^1\frac{dx}{x+1}=$$

$$=\ln 2-1+\ln(1+x)\big|_0^1=-1+2\ln 2.$$

THÉORÈME 3 (DE LA MOYENNE POUR UNE INTÉGRALE DÉFINIE). *Pour toute fonction $f$ continue sur un intervalle $[a,\ b]$ il existe un point $\xi\in\,]a,\ b[$ tel que*

$$\int_a^b f(x)\,dx=f(\xi)\,(b-a). \tag{13}$$

DÉMONSTRATION. La fonction $f$ étant continue, elle possède une primitive $\Phi$ et par suite

$$\int_a^b f(x)\,dx=\Phi(b)-\Phi(a)=\Phi'(\xi)\,(b-a)=f(\xi)\,(b-a),\qquad \xi\in\,]a,\ b[. \tag{14}$$

La première égalité de (14) est la formule de Newton-Leibniz pour une fonction $f$ continue sur $[a,\ b]$. La seconde est la formule des accroissements finis pour $\Phi$. La troisième enfin résulte de ce que $\Phi'(x)=f(x),\ \ \forall\, x\in[a,\ b]$.

## § 5. Reste de la formule de Taylor sous la forme intégrale

Supposons qu'une fonction $f(x)$ est dérivable jusqu'à l'ordre $n+1$ compris. Alors, en vertu de la formule de Newton-Leibniz, on a

$$f(x)=f(a)+\int_a^x f'(t)\,dt=\left(\begin{matrix}u=f'(t)\\ dv=dt\end{matrix}\ \middle|\ \begin{matrix}du=f''(t)\,dt\\ v=t-x\end{matrix}\right)=$$

$$= f(a) + (t-x) f'(t) \Big|_{t=a}^{t=x} - \int_a^x (t-x) f''(t)\, dt =$$

$$= f(a) + f'(a)(x-a) + \int_a^x (x-t) f''(t)\, dt =$$

$$= \left( \begin{array}{l|l} u = f''(t) & du = f'''(t)\, dt \\ (x-t)\, dt = dv & v = -\frac{(x-t)^2}{2!} \end{array} \right) =$$

$$= f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \int_a^x f'''(t) \frac{(x-t)^2}{2!}\, dt.$$

En poursuivant ce processus d'intégration par parties, on obtient

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + r_n(x), \tag{1}$$

où

$$r_n(x) = \frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t)\, dt. \tag{2}$$

La formule (1) s'appelle *formule de Taylor avec un reste sous la forme intégrale* (2).

En appliquant le théorème 3 de la moyenne (§ 4) à l'intégrale (2), on obtient

$$r_n(x) = \frac{1}{n!}(x-\xi)^n f^{(n+1)}(\xi)(x-a), \quad \xi \in ]a, x[.$$

En posant

$$\xi = a + \theta(x-a), \quad 0 < \theta < 1,$$

on obtient

$$r(x) = \frac{(x-a)^{n+1}}{n!}(1-\theta)^n f^{(n+1)}(a + \theta(x-a)),$$

c'est-à-dire le reste de Cauchy de la formule de Taylor en $a$ (cf. chap. 4, § 14, (10)).

## § 6. Sommes de Darboux [1]. Conditions d'existence de l'intégrale

Soit donnée une fonction $f(x)$ bornée sur un intervalle $[a, b]$ : ($|f(x)| \leqslant M$). Considérons une subdivision de $[a, b]$

$$R: a = x_0 < x_1 < \ldots < x_n = b$$

et posons

$$m_i = \inf_{x \in [x_i, x_{i+1}]} f(x), \quad M_i = \sup_{x \in [x_i, x_{i+1}]} f(x).$$

---

[1]) Gaston Darboux, mathématicien français (1842-1917).

Outre les sommes de Riemann

$$\sigma_R = \sum_{i=0}^{n-1} f(\xi_i)\,\Delta x_i$$

considérons les sommes

$$s_R = \sum_{i=0}^{n-1} m_i \Delta x_i, \quad S_R = \sum_{i=0}^{n-1} M_i \Delta x_i,$$

dites *sommes inférieure et supérieure de Darboux.* Il est évident que $s_R \leqslant S_R$.

Les sommes de Darboux ne sont pas nécessairement des sommes de Riemann. Mais si $f(x)$ est une fonction continue, alors $s_R$ et $S_R$ sont respectivement la plus petite et la plus grande des sommes de Riemann attachées à une subdivision donnée, puisque le théorème de Weierstrass nous dit que la fonction $f(x)$ présente son minimum et son maximum dans chaque intervalle $[x_i, x_{i+1}]$, donc on peut exhiber des points $\xi_i, \xi_i' \in [x_i, x_{i+1}]$ tels que $f(\xi_i) = m_i$ et $f(\xi_i') = M_i$.

Comme $m_i \leqslant f(x) \leqslant M_i$ et que $\Delta x_i > 0$, il vient

$$s_R \leqslant \sigma_R \leqslant S_R. \tag{1}$$

Si l'on fige la subdivision $R$, les sommes $s_R$ et $S_R$ sont des nombres constants, mais $\sigma_R$ est variable, car les $\xi_i$ sont arbitraires. Il est immédiat de voir qu'en choisissant convenablement les points $\xi_i$ on peut rendre la somme $\sigma_R$ aussi proche que l'on veut de $s_R$ et $S_R$, autrement dit $s_R$ et $S_R$ sont les bornes inférieure et supérieure des sommes de Riemann attachées à la subdivision $R$:

$$s_R = \sum_{i=0}^{n-1} m_i \Delta x_i = \inf_{\xi_i} \sum_{i=0}^{n-1} f(\xi_i)\,\Delta x_i,$$

$$S_R = \sum_{i=0}^{n-1} M_i \Delta x_i = \sup_{\xi_i} \sum_{i=0}^{n-1} f(\xi_i)\,\Delta x_i.$$

Soient $R_1$, $R_2$ et $R_3$ des subdivisions de $[a, b]$. Si tous les points de $R_1$ appartiennent à $R_2$, on dira que $R_2$ est un prolongement de $R_1$ et on écrira $R_1 \subset R_2$. Si $R_3$ est composée des points de $R_1$ et des points de $R_2$, on écrira $R_3 = R_1 + R_2$.

Propriétés des sommes de Darboux:

1° *Si l'on ajoute des points à une subdivision $R$, la somme supérieure de Darboux décroîtra et la somme inférieure croîtra:*

$$S_{R'} \leqslant S_R, \quad s_R \leqslant s_{R'}, \quad \forall R \subset R'.$$

Donc,

$$S_{R'} - s_{R'} \leqslant S_R - s_R.$$

DÉMONSTRATION. On peut de toute évidence se limiter au cas où l'on ajoute un seul point $x' \in \,]x_i, x_{i+1}[$. Soient $S_R$ et $S_{R'}$ les sommes supérieures de Darboux correspondant respectivement à $R$ et à $R'$. Alors la somme $S_{R'}$ se distingue de $S_R$ par le fait qu'à la place du terme $M_i\,\Delta x_i$ elle contient deux termes:

$$M_i'(x' - x_i) \quad \text{et} \quad M_i''(x_{i+1} - x'),$$

où $M_i' = \sup\limits_{x \in [x_i, x']} f(x)$ et $M_i'' = \sup\limits_{x \in [x', x_{i+1}]} f(x)$. Les intervalles $[x_i, x']$ et $[x', x_{i+1}]$ étant des parties de $[x_i, x_{i+1}]$, on a $M_i' \leqslant M_i$, $M_i'' \leqslant M_i$ (lorsque le domaine d'étude se rétrécit, sup ne peut que diminuer). Donc,

$$M_i'(x' - x_i) + M_i''(x_{i+1} - x') \leqslant M_i(x' - x_i + x_{i+1} - x') = M_i(x_{i+1} - x_i)$$

autrement dit $S_{R'} \leqslant S_R$. C.q.f.d.

La démonstration est identique pour les sommes inférieures.

2° *Toute somme inférieure de Darboux est inférieure à toute somme supérieure de Darboux, même si elle correspond à une autre subdivision de l'intervalle*: $s_{R_1} \leqslant S_{R_2}$.

DEMONSTRATION. Soient $R_3 = R_1 + R_2$. La propriété 1° nous donne $s_{R_1} \leqslant s_{R_3} \leqslant S_{R_3} \leqslant S_{R_2}$.

Nous avons ainsi démontré que l'ensemble $\{s_R\}$ des sommes inférieures de Darboux est majoré par une somme supérieure $S_{R'}$ ($s_R \leqslant S_{R'}$), donc cet ensemble possède une borne supérieure:

$$I_* = \sup_R s_R \leqslant S_{R'}.$$

Nous avons prouvé en même temps que toute somme supérieure $S_{R'}$ est supérieure à $I_*$. Ceci montre que *l'ensemble des sommes supérieures admet une borne inférieure*

$$I^* = \inf_{R'} S_{R'} \geqslant I_*.$$

Ainsi, $I_* \leqslant I^*$. De plus, pour toute subdivision $R$

$$s_R \leqslant I_* \leqslant I^* \leqslant S_R. \tag{2}$$

Les nombres $I_*$ et $I^*$ s'appellent respectivement *intégrales inférieure et supérieure de Darboux.*

THEOREME (D'EXISTENCE DE L'INTÉGRALE). *Pour que l'intégrale définie d'une fonction bornée $f(x)$ existe, il est nécessaire et suffisant que*

$$\lim_{\lambda_R \to 0} (S_R - s_R) = \lim_{\lambda_R \to 0} \sum_{i=0}^{n-1} \omega_i \Delta x_i = 0, \tag{3}$$

*où le nombre $\omega_i = M_i - m_i$ est l'étendue de variation de la fonction $f(x)$ sur $[x_i, x_{i+1}]$.*

DEMONSTRATION. *Condition nécessaire.* Supposons que l'intégrale définie $I$ de la fonction $f(x)$ existe, c'est-à-dire que pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que pour $\lambda_R < \delta$ on a $I - \varepsilon < \sigma_R < I + \varepsilon$, quel que soit $\xi_i \in [x_i, x_{i+1}]$.

Nous avons vu plus haut que $s_R$ et $S_R$ étaient les bornes inférieure et supérieure des sommes de Riemann $\sigma_R$ attachées à $R$. Donc

$$I - \varepsilon \leqslant s_R \leqslant \sigma_R \leqslant S_R \leqslant I + \varepsilon, \quad \forall R \quad \lambda_R < \delta,$$

i.e.

$$\lim_{\lambda_R \to 0} s_R = I, \quad \lim_{\lambda_R \to 0} S_R = I$$

et

$$\lim_{\lambda_R \to 0} (S_R - s_R) = 0.$$

*Condition suffisante.* Supposons que la condition (3) est réalisée. De l'inégalité (2) il s'ensuit alors que $I_* = I^*$. Posons $I = I_* = I^*$. Alors

$$s_R \leqslant I \leqslant S_R. \tag{4}$$

De (3) il résulte que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que $|S_R - s_R| < \varepsilon$ pour $\lambda_R < \delta$. De (1) et (4) on déduit alors que

$$|I - \sigma_R| < \varepsilon \quad \text{pour} \quad \lambda_R < \delta,$$

c'est-à-dire que $I$ est la limite de $\sigma_R$ et $f(x)$ est intégrable.

REMARQUE. De la démonstration du théorème, il résulte que si la fonction $f(x)$ est intégrable sur $[a, b]$, alors $\lim_{\lambda_R \to 0} s_R = \lim_{\lambda_R \to 0} S_R = \int_a^b f(x)\,dx$, et réciproquement si $\lim_{\lambda_R \to 0} s_R = \lim_{\lambda_R \to 0} S_R = I$, alors $I = \int_a^b f(x)\,dx$.

## § 7. Intégrabilité des fonctions continues et monotones

THÉORÈME 1. *Si une fonction $f(x)$ est continue sur un intervalle $[a, b]$, alors elle est intégrable sur cet intervalle.*

DÉMONSTRATION. La fonction $f(x)$ est continue sur $[a, b]$, donc uniformément continue, et par suite pour tout $\varepsilon > 0$ il existe un $\delta(\varepsilon) > 0$ tel que pour toutes les subdivisions de $[a, b]$ de pas $\lambda_R < \delta$ l'on ait $\omega_i < \varepsilon$. D'où

$$\sum_{i=0}^{n-1} \omega_i \Delta x_i \leqslant \varepsilon \sum_{i=0}^{n-1} \Delta x_i = \varepsilon(b-a).$$

Comme $\varepsilon$ est arbitraire, on conclut que $\lim_{\lambda_R \to 0} \sum_{i=0}^{n-1} \omega_i \Delta x_i = 0$, et le théorème du § 6 nous dit que la fonction $f(x)$ est intégrable.

THÉORÈME 2. *Une fonction monotone sur un intervalle fermé est intégrable sur cet intervalle.*

DÉMONSTRATION. Pour fixer les idées nous admettrons que la fonction $f(x)$ est croissante. Nous admettrons de même que $f(a) < f(b)$, sinon la fonction serait constante et le théorème trivial.

Comme $f(a) \leqslant f(x) \leqslant f(b)$, $\forall x \in [a, b]$, la fonction $f$ est bornée sur $[a, b]$. Soit une subdivision $R$ de $[a, b]$ telle que $\lambda_R < \delta$. Puisque $\omega_i = f(x_{i+1}) - f(x_i)$, on a

$$\sum_{i=0}^{n-1} \omega_i \Delta x_i \leqslant \delta \sum_{i=0}^{n-1} \omega_i = \delta [f(x_1) - f(x_0) + f(x_2) - f(x_1) + \ldots + f(x_n) - f(x_{n-1})] = \delta [f(b) - f(a)],$$

$x_0 = a$, $x_n = b$. Si maintenant l'on prend $\delta = \varepsilon/[f(b) - f(a)]$, on obtient

$$\sum_{i=0}^{n-1} \omega_i \Delta x_i \leqslant \varepsilon,$$

et le théorème d'existence (théorème du § 6) nous dit que $f(x)$ est intégrable. C.q.f.d.

REMARQUE 1. Signalons qu'une fonction monotone peut présenter un ensemble dénombrable de points de discontinuité. Par exemple, la fonction $y = \left\{ x + \frac{1}{n}, \frac{1}{n+1} < x \leqslant \frac{1}{n}, \ n = 1, 2, \ldots \right\}$ est monotone croissante sur $[0, 1]$ et possède un ensemble dénombrable de points de discontinuité. Donc, elle est intégrable d'après le théorème 2.

REMARQUE 2. Si $f(x)$ est intégrable sur $[a, b]$, il en est de même de $|f(x)|$. En effet, pour tous $x'$ et $x''$ de $[x_i, x_{i+1}]$ on a

$$||f(x')| - |f(x'')|| \leqslant |f(x') - f(x'')|. \quad (1)$$

Si $\omega_i^*$ et $\omega_i$ sont les étendues de variation respectives de $|f(x)|$ et de $f(x)$ sur $[x_i, x_{i+1}]$, alors de (1) il résulte que $\omega_i^* \leqslant \omega_i$ et

$$\sum_{i=0}^{n-1} \omega_i^* \Delta x_i \leqslant \sum_{i=0}^{n-1} \omega_i \Delta x_i.$$

Comme $f(x)$ est intégrable, on a

$$\sum_{i=0}^{n-1} \omega_i \Delta x_i \to 0 \text{ pour } \lambda_R \to 0,$$

donc

$$\sum_{i=0}^{n-1} \omega_i^* \Delta x_i \to 0,$$

et par suite, $|f(x)|$ est intégrable.

## § 8. Intégrales impropres

Soit donnée une fonction $f$ sur un intervalle semi-ouvert borné $[a, b[$. Supposons que cette fonction est intégrable (par exemple, est continue ou continue par morceaux) sur tout intervalle fermé $[a, b']$, où $b' < b$ et qu'elle n'est pas bornée au voisinage de $b$. Son intégrale au sens de Riemann sur $[a, b[$, ou ce qui revient au même sur $[a, b]$, n'existe pas, car toute fonction intégrable-Riemann sur $[a, b]$ est nécessairement bornée. Néanmoins, il est possible qu'existe la limite finie

$$\lim_{b' \to b} \int_a^{b'} f(x)\, dx.$$

Si tel est le cas, cette limite s'appelle *intégrale impropre de $f$ sur* $[a, b[$ et se note

$$\int_a^b f(x)\, dx = \lim_{b' \to b} \int_a^{b'} f(x)\, dx. \quad (1)$$

On dit alors que *l'intégrale* $\int_a^b f(x)\, dx$ *est convergente.* Dans le cas contraire, elle *diverge* ou n'existe pas en tant qu'intégrale impropre.

Supposons maintenant que la fonction $f$ est définie sur la section $[a, \infty[$ et intégrable sur tout intervalle fermé borné $[a, b']$, où

$a < b' < \infty$. Si la limite

$$\lim_{b' \to \infty} \int_a^{b'} f(x)\,dx$$

existe, on l'appelle *intégrale impropre de f sur* $[a, \infty[$ et on la note:

$$\int_a^{\infty} f(x)\,dx = \lim_{b' \to \infty} \int_a^{b'} f(x)\,dx.$$

Convenons de la terminologie suivante. L'expression

$$\int_a^b f(x)\,dx \qquad (2)$$

sera appelée *intégrale de f à singularité unique en b* si sont remplies les conditions suivantes: si $b$ est un point fini, alors la fonction $f$ est intégrable sur $[a, b']$, quel que soit $b'$ tel que $a < b' < b$, et de plus n'est pas bornée au voisinage de $b$. Si $b = +\infty$, on admet seulement que la fonction $f$ est intégrable sur $[a, b']$ pour tout $b' > a$ fini.

On définit de façon analogue l'intégrale $\int_a^b f(x)\,dx$ à singularité unique en $a$. Supposons maintenant que $b$ est un point fini. Si $a < b$ est aussi un point fini, alors au voisinage de $a$ la fonction $f$ n'est pas bornée et est intégrable sur tout intervalle $[a', b]$, où $a < a' < b$. Si $a = -\infty$, on admet que la fonction $f$ est intégrable sur $[a', b]$ pour tout $a' < b$.

Dans la suite, on étudiera, pour fixer les idées, une intégrale (2) à singularité unique en $b$ fini ou infini. Tous les résultats peuvent être généralisés au cas d'une intégrale à singularité unique en $a$.

THÉORÈME. *Soit donnée une intégrale* (2) *à singularité unique en b. Pour que cette intégrale converge il est nécessaire et suffisant que soit remplie la condition (de Cauchy): pour tout* $\varepsilon > 0$ *il existe un* $b_0 < b$ *tel que*

$$\left| \int_{b'}^{b''} f(t)\,dt \right| < \varepsilon, \qquad (3)$$

*quels que soient* $b'$ *et* $b''$ *tels que* $b_0 < b' < b'' < b$.

DÉMONSTRATION. Considérons la fonction

$$F(x) = \int_a^x f(t)\,dt \quad (a < x < b).$$

Dire que l'intégrale (2) converge revient à dire que la limite $\lim\limits_{\substack{x\to b\\ x<b}} F(x)$ existe ou ce qui revient au même qu'est remplie la condition de Cauchy: pour tout $\varepsilon > 0$ on peut exhiber un $b_0 \in ]a, b[$ tel que $|F(b'') - F(b')| < \varepsilon$ pour tous les $b'$ et $b''$ vérifiant les inégalités $b_0 < b' < b'' < b$. Or,

$$F(b'') - F(b') = \int_{b'}^{b''} f(t)\,dt,$$

ce qui prouve le théorème.

Exemple 1. L'intégrale

$$\int_0^1 \frac{dx}{x^\alpha}, \tag{4}$$

où $\alpha > 0$ est une constante, présente une seule singularité en $x = 0$. Pour étudier sa convergence il faut calculer la limite

$$\lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}} \int_\varepsilon^1 \frac{dx}{x^\alpha} = \lim_{\varepsilon\to 0} \frac{x^{1-\alpha}}{1-\alpha}\Big|_\varepsilon^1 = \lim_{\varepsilon\to 0} \frac{1}{1-\alpha}[1-\varepsilon^{1-\alpha}] = \begin{cases} \dfrac{1}{1-\alpha}, & \alpha<1,\\ \infty, & \alpha>1. \end{cases}$$

Donc, l'intégrale (4) converge vers $(1-\alpha)^{-1}$ pour $\alpha < 1$ et diverge pour $\alpha > 1$. Si $\alpha = 1$, elle diverge:

$$\lim_{\varepsilon\to 0} \int_\varepsilon^1 \frac{dx}{x} = -\lim_{\varepsilon\to 0} \ln \varepsilon = +\infty.$$

Exemple 2.

$$\int_1^\infty \frac{dx}{x^\alpha} = \lim_{N\to\infty} \int_1^N \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} \lim_{N\to\infty} x^{1-\alpha}\,\Big|_1^N = \begin{cases} \dfrac{1}{\alpha-1}, & \alpha>1 \text{ (converge)},\\ +\infty, & \alpha<1 \text{ (diverge)}, \end{cases}$$

$$\int_1^\infty \frac{dx}{x} = \lim_{N\to\infty} \int_1^N \frac{dx}{x} = \lim_{N\to\infty} \ln N = +\infty \text{ (diverge)}.$$

Exemple 3. L'intégrale $\int\limits_0^\infty e^{-x}\,dx$ possède une seule singularité au point $x = +\infty$. Elle converge vers 1:

$$\int_0^\infty e^{-x}\,dx = \lim_{N\to+\infty} \int_0^N e^{-x}\,dx = \lim_{N\to+\infty} (-e^{-x})\,\Big|_0^N = \lim_{N\to+\infty} [1-e^{-N}] = 1.$$

Considérons de nouveau l'intégrale

$$\int_a^b f(x)\,dx \tag{5}$$

à singularité unique en $b$. Alors l'intégrale

$$\int_c^b f(x)\,dx, \tag{6}$$

où $a < c < b$ présente aussi une seule singularité en $b$. La condition de Cauchy se formule de façon analogue pour les intégrales (5) et (6). Donc, ces dernières convergent ou divergent simultanément. De plus pour $a < c < b$ on a de toute évidence

$$\int_a^b f\,dx = \lim_{b'\to b} \int_a^{b'} f\,dx = \lim_{b'\to b} \left( \int_a^c f\,dx + \int_c^{b'} f\,dx \right) =$$
$$= \int_a^c f\,dx + \lim_{b'\to b} \int_c^{b'} f\,dx = \int_a^c f\,dx + \int_c^b f\,dx, \tag{7}$$

où $\int_a^c$ est une intégrale de Riemann ordinaire, $\int_a^b$ et $\int_c^b$ des intégrales impropres.

Signalons la relation

$$\int_a^b (Af + B\varphi)\,dx = \lim_{b'\to b} \int_a^{b'} (Af + B\varphi)\,dx =$$
$$= A \lim_{b'\to b} \int_a^{b'} f\,dx + B \lim_{b'\to b} \int_a^{b'} \varphi\,dx = A \int_a^b f\,dx + B \int_a^b \varphi\,dx, \tag{8}$$

où $A$ et $B$ sont des constantes. Cette relation exprime le fait suivant: si les intégrales du second membre existent, il en est de même de l'intégrale du premier, et l'on a le signe d'égalité.

On dit que l'intégrale (5) *est absolument convergente* si l'intégrale

$$\int_a^b |f(x)|\,dx \tag{9}$$

converge.

*Une intégrale absolument convergente est convergente.* En effet, de la convergence de l'intégrale (9) il s'ensuit que pour tout $\varepsilon > 0$

on peut exhiber un point $b_0 \in ]a, b[$, tel que

$$\varepsilon > \int_{b'}^{b''} |f(x)|\, dx \geqslant \left| \int_{b'}^{b''} f(x)\, dx \right|$$

pour $b_0 < b' < b'' < b$, autrement dit la condition de Cauchy est remplie pour l'intégrale (5). Comme

$$\left| \int_a^{b'} f(x)\, dx \right| \leqslant \int_a^{b'} |f(x)|\, dx,$$

par passage à la limite pour $b' \to b$, on obtient

$$\left| \int_a^{b} f(x)\, dx \right| \leqslant \int_a^{b} |f(x)|\, dx. \qquad (10)$$

REMARQUE. L'inégalité (10) est valable également pour une intégrale non absolument convergente: dans ce cas au second membre figure $\infty$. Ceci est largement utilisé dans les calculs. Si l'on a à étudier la convergence de l'intégrale $\int_a^b f\, dx$, on écrit l'inégalité (10) et on étudie la convergence de l'intégrale $\int_a^b |f|\, dx$. Si cette dernière converge, c'est-à-dire si $\int_a^b |f|\, dx < \infty$, alors il en est de même de l'intégrale $\int_a^b f\, dx$. Si $\int_a^b |f|\, dx = \infty$, il faudra faire appel à des méthodes plus subtiles. Il est possible que cette intégrale converge, mais surtout pas absolument (voir exemples à la fin du § 9).

## § 9. Intégrales impropres de fonctions à valeurs réelles positives

Soit donnée l'intégrale

$$\int_a^b f(x)\, dx \qquad (1)$$

à une seule singularité en $b$ d'une fonction $f(x) \geqslant 0$ sur l'intervalle d'intégration $[a, b[$. Alors la fonction

$$F(b') = \int_a^{b'} f(x)\,dx \quad (a < b' < b)$$

est de toute évidence monotone croissante. Donc, si elle est majorée par un nombre $M$, alors l'intégrale (1) converge et

$$\int_a^b f(x)\,dx = \lim_{b' \to b} \int_a^{b'} f(x)\,dx \leqslant M.$$

Si $F$ n'est pas majorée, alors l'intégrale (1) diverge:

$$\int_a^b f(x)\,dx = \lim_{b' \to b} \int_a^{b'} f(x)\,dx = +\infty.$$

Si $f(x) \geqslant 0$ sur $[a, b[$, on écrit

$$\int_a^b f(x)\,dx < \infty \quad \text{ou} \quad \int_a^b f(x)\,dx = \infty$$

selon que l'intégrale converge ou diverge.

Théorème 1. *Supposons que les intégrales*

$$\int_a^b f(x)\,dx, \tag{1}$$

$$\int_a^b \varphi(x)\,dx \tag{2}$$

*présentent une seule singularité en $b$ et que*

$$0 \leqslant f(x) \leqslant \varphi(x) \tag{3}$$

*sur* $[a, b[$.

*Si l'intégrale* (2) *converge, il en est de même de l'intégrale* (1), *et l'on a*

$$\int_a^b f\,dx \leqslant \int_a^b \varphi\,dx$$

*Si l'intégrale* (1) *diverge, il en est de même de l'intégrale* (2).

Démonstration. De (3) il s'ensuit que

$$\int_a^{b'} f\,dx \leqslant \int_a^{b'} \varphi\,dx, \tag{4}$$

où $b' \in ]a, b[$. Si l'intégrale (2) converge, le second membre de (4) est majoré par un nombre égal à l'intégrale (2). Donc, le premier membre est aussi majoré par ce nombre. Le premier membre étant une fonction monotone croissante de $b'$, il converge vers l'intégrale:

$$\int_a^b f\,dx = \lim_{b'\to b}\int_a^{b'} f\,dx \leqslant \int_a^b \varphi\,dx.$$

Si l'intégrale (1) diverge, le premier membre de (4) tend vers $\infty$ pour $b' \to b$, donc le second tend aussi vers $\infty$.

THÉORÈME 2. *Supposons que les intégrales* (1) *et* (2) *présentent une seule singularité en* $b$, *que les intégrants sont des fonctions positives et qu'existe la limite*

$$\lim_{x\to b}\frac{f(x)}{\varphi(x)} = A > 0. \tag{5}$$

*Sous ces conditions, ces intégrales divergent ou convergent toutes deux.*

DÉMONSTRATION. De (5) il s'ensuit que pour tout nombre strictement positif $\varepsilon < A$ on peut exhiber un $c \in [a, b[$ tel que

$$A-\varepsilon < \frac{f(x)}{\varphi(x)} < A+\varepsilon \quad (c < x < b),$$

et comme $\varphi(x) > 0$, alors

$$(A-\varepsilon)\,\varphi(x) < f(x) < (A+\varepsilon)\,\varphi(x), \quad x \in ]c, b[. \tag{6}$$

La convergence de l'intégrale $\int_a^b \varphi\,dx$ implique celle de $\int_c^b \varphi\,dx$ et de $\int_c^b (A+\varepsilon)\,\varphi\,dx$; donc, le théorème précédent nous dit que l'intégrale $\int_c^b f\,dx$ converge aussi et avec elle $\int_a^b f\,dx$. Réciproquement, la convergence de $\int_a^b f\,dx$ entraîne celle de $\int_c^b \varphi\,dx$, car, outre (5), on a

$$\lim_{x\to b}\frac{\varphi(x)}{f(x)} = \frac{1}{A} > 0.$$

REMARQUE. La relation (5) exprime que la fonction $f$ est équivalente à la fonction $A\varphi$ pour $x \to b$.

EXEMPLE 1. Etudier la convergence de l'intégrale

$$\int_0^\infty \sin kxe^{-x}\,dx.$$

On a

$$\left|\int_0^\infty e^{-x}\sin kx\,dx\right| \leqslant \int_0^\infty |e^{-x}\sin kx|\,dx \leqslant \int_0^\infty e^{-x}\,dx = 1 < \infty.$$

On s'est servi de l'inégalité (10) du § 8 et de la remarque qui la suit.

Le signe $\sim$ traduira le fait que les intégrales convergeront ou divergeront toutes deux en vertu du théorème 2.

EXEMPLE 2.

$$\int_0^1 \frac{dx}{\sin x} \sim \int_0^1 \frac{dx}{x} = \infty.$$

EXEMPLE 3.

$$\int_0^1 \frac{dx}{\sin\sqrt{x}} \sim \int_0^1 \frac{dx}{\sqrt{x}} < \infty.$$

EXEMPLE 4.

$$\int_1^\infty \frac{x-1}{x}e^{-x}\,dx \sim \int_1^\infty e^{-x}\,dx < \infty.$$

Les intégrales des exemples 2 et 3 présentent une seule singularité en $x=0$. D'autre part, $\sin x \approx x$, $\sin\sqrt{x} \approx \sqrt{x}$, $x \to 0$.

L'intégrale de l'exempe 4 présente une seule singularité en $x=\infty$. De plus, $\frac{x-1}{x}e^{-x} \approx e^{-x}$, $x \to +\infty$.

EXEMPLE 5. $\int_0^\infty (x^2-3x+5)\,e^{-x}\,dx$ converge, car

$$\left|\int_0^\infty (x^2-3x+5)\,e^{-x}\,dx\right| \leqslant \int_0^\infty |(x^2-3x+5)\,e^{-x/2}|\,e^{-x/2}\,dx \leqslant$$

$$\leqslant M\int_0^\infty e^{-x/2}\,dx < \infty.$$

En effet, $\lim\limits_{x\to\infty}(x^2-3x+5)\,e^{-x/2}=0$, donc il existe un $N>0$ tel que $|(x^2-3x+5)\,e^{-x/2}|<1$, $\forall\; x>N$.

Par ailleurs, étant continue sur $[0, N]$ la fonction $|(x^2 - 3x + 5)e^{-x/2}|$ est majorée par un nombre $M_1$. Donc, elle est majorée sur $[0, \infty[$ par le nombre $M = \max\{1, M_1\}$.

## § 10. Intégration par parties d'intégrales impropres

EXEMPLE 1. Les intégrales impropres

$$\int_a^\infty \frac{\sin x}{x}\,dx, \quad \int_a^\infty \frac{\cos x}{x}\,dx \quad (a>0) \tag{1}$$

convergent. En effet, en intégrant par parties, on obtient

$$\int_a^A \frac{\sin x}{x}\,dx = -\frac{\cos x}{x}\Big|_a^A - \int_a^A \frac{\cos x}{x^2}\,dx$$

pour tous les $A>a$ finis. En passant à la limite pour $A \to \infty$, on trouve

$$\int_a^\infty \frac{\sin x}{x}\,dx = \frac{\cos a}{a} - \int_a^\infty \frac{\cos x}{x^2}\,dx,$$

où l'intégrale du second membre converge absolument :

$$\left|\int_a^\infty \frac{\cos x}{x^2}\,dx\right| \leqslant \int_a^\infty \frac{|\cos x|}{x^2}\,dx \leqslant \int_a^\infty \frac{dx}{x^2} = \frac{1}{a} < \infty.$$

EXEMPLE 2. L'intégrale $\int_a^\infty \frac{\sin x}{x}\,dx$ ne converge pas absolument (est semi-convergente), car l'intégrale

$$\int_a^\infty \frac{|\sin x|}{x}\,dx = \infty, \quad a>0. \tag{2}$$

En effet, en vertu de l'inégalité $\sin^2 x \leqslant |\sin x|$, l'intégrale impropre :

$$\int_a^\infty \frac{|\sin x|}{x}\,dx \geqslant \int_a^\infty \frac{\sin^2 x}{x}\,dx = \int_a^\infty \frac{1-\cos 2x}{2x}\,dx =$$

$$= (2x = u) = \frac{1}{2}\int_{2a}^\infty \frac{1-\cos u}{u}\,du = \frac{1}{2}\int_{2a}^\infty \frac{du}{u} - \frac{1}{2}\int \frac{\cos u}{u}\,du.$$

Or, $\int\limits_{2a}^{\infty} u^{-1} \cos u \, du$ converge et $\int\limits_{2a}^{\infty} u^{-1} \, du$ diverge, donc l'intégrale impropre (2) diverge.

REMARQUE. L'intégrale (1) converge, car la fonction $\sin x$ est périodique et prend des valeurs tour à tour positives et négatives qui se compensent.

Ce phénomène sera justifié dans la théorie des séries (voir série de Leibniz et séries semi-convergentes).

Les exemples cités montrent que l'intégration par parties est parfois un instrument efficace d'étude de la convergence des intégrales impropres.

Les considérations générales qui vont suivre nous permettront de mieux comprendre le mécanisme de cette méthode.

Soient $\varphi(x)$ une fonction continue sur $[a, \infty[$, $\Phi(x)$ une de ses primitives. On suppose d'autre part que $g(x)$ est une fonction continûment dérivable sur $[a, \infty[$. Alors

$$\int\limits_a^A \varphi(x) g(x) \, dx = g(x) \Phi(x) \Big|_a^A - \int\limits_a^A \Phi(x) g'(x) \, dx =$$

$$= g(A) \Phi(A) - g(a) \Phi(a) - \int\limits_a^A \Phi(x) g'(x) \, dx. \qquad (3)$$

Si

1) $\lim\limits_{A \to \infty} g(A) \Phi(A) = 0,$

2) l'intégrale $\int\limits_a^{\infty} \Phi(x) g'(x) \, dx$ est convergente, il est alors évident que l'intégrale impropre

$$\int\limits_a^{\infty} \varphi(x) g(x) \, dx = \lim_{A \to \infty} \int\limits_a^A \varphi(x) g(x) \, dx = -g(a) \Phi(a) - \int\limits_a^{\infty} \Phi(x) g'(x) \, dx \qquad (4)$$

est convergente.

De là il résulte notamment le

CRITÈRE DE DIRICHLET DE CONVERGENCE DE L'INTÉGRALE (4). Si une fonction $\Phi(x)$ est majorée par un nombre $M$, et $g(x)$ est décroissante et converge vers 0 pour $x \to \infty$, alors l'intégrale (4) est convergente.

Il est immédiat que ces conditions entraînent la propriété 1). D'autre part

$$\left| \int\limits_a^{\infty} \Phi(x) g'(x) \, dx \right| \leqslant \int\limits_a^{\infty} | \Phi(x) g'(x) | \, dx \leqslant M \int\limits_a^{\infty} | g'(x) | \, dx =$$

$$= -M \int\limits_a^{\infty} g'(x) \, dx = -M \lim_{A \to \infty} \int\limits_a^A g'(x) \, dx = -M \lim_{A \to \infty} [g(A) - g(a)] = g(a) \cdot M.$$

EXEMPLE 3. L'intégrale

$$\int_a^{\infty} \frac{\sin x}{x^{\alpha}}\, dx \quad (\alpha > 0)$$

présente une seule singularité en $x = \infty$, converge pour $\alpha > 0$. Ceci résulte du critère de Dirichlet dans lequel on pose $g(x) = x^{-\alpha}$, $\varphi(x) = \sin x$, $\Phi(x) = -\cos x$ ($|\Phi(x)| \leqslant 1$). La convergence absolue n'a lieu que pour $\alpha > 1$ et se démontre comme dans l'exemple 2 du § 9.

## § 11. Intégrale impropre à singularités en plusieurs points

Soit donnée l'intégrale

$$\int_a^b f(x)\, dx, \tag{1}$$

où $f(x)$ est une fonction définie sur $]a, b[$, $a$ et $b$ pouvant être finis ou infinis.

Supposons que l'intervalle $]a, b[$ peut être partagé en un nombre fini d'intervalles par les points $a = c_0 < c_1 < \ldots < c_N = b$ de telle sorte que chaque intégrale

$$\int_{c_k}^{c_{k+1}} f(x)\, dx \quad (k = 0, 1, \ldots, N-1) \tag{2}$$

présente une seule singularité soit en $c_k$, soit en $c_{k+1}$.

Si toutes les intégrales impropres (2) sont convergentes (resp. absolument convergentes), on dit que l'intégrale (1) est *convergente* (resp. *absolument convergente*) et on lui affecte le nombre

$$\int_a^b f(x)\, dx = \sum_{k=0}^{N-1} \int_{c_k}^{c_{k+1}} f(x)\, dx.$$

Si l'une au moins des intégrales (2) est divergente, on admet que l'intégrale (1) est divergente.

Si $f(x) \geqslant 0$, par analogie avec les intégrales à une singularité, on conviendra d'écrire

$$\int_a^b f(x)\, dx < \infty$$

si (1) est convergente, et

$$\int_a^b f(x)\, dx = \infty$$

si elle est divergente.

EXEMPLE 1.

$$\int_{-\infty}^{\infty} e^{-x}\,dx = \int_{-\infty}^{0} e^{-x}\,dx + \int_{0}^{\infty} e^{-x}\,dx = \infty + 1 = \infty.$$

Cette intégrale présente deux singularités, une en $x = -\infty$, l'autre en $x = +\infty$. Aussi l'avons-nous représentée par une somme de deux intégrales présentant chacune une singularité. De toute évidence

$$\int_{-\infty}^{0} e^{-x}\,dx = \infty, \quad \int_{0}^{\infty} e^{-x}\,dx = 1.$$

Nous avons admis que $\infty \pm 1 = \infty$.

EXEMPLE 2. $(\alpha > 0)$

$$\int_{0}^{\infty} \frac{\sin x}{x^{\alpha}}\,dx \begin{cases} \text{est semi-convergente pour } 0 < \alpha \leqslant 1, \\ \text{absolument convergente pour } 1 < \alpha < 2, \\ \text{divergente pour } \alpha \geqslant 2. \end{cases} \tag{3}$$

En effet, cette intégrale a deux singularités en $x = 0$ et $x = \infty$, donc on peut la représenter par la somme

$$\int_{0}^{\infty} \frac{\sin x}{x^{\alpha}}\,dx = \int_{0}^{1} \frac{\sin x}{x^{\alpha}}\,dx + \int_{1}^{\infty} \frac{\sin x}{x^{\alpha}}\,dx.$$

L'intégrant de la première intégrale étant une fonction strictement positive, cette intégrale est soit divergente, soit absolument convergente. Pour l'étudier nous aurons besoin des inégalités suivantes (cf. chap. 3, § 3, (6) et chap. 4, § 9, exemple 1)

$$\frac{2}{\pi} x^{1-\alpha} \leqslant \frac{\sin x}{x^{\alpha}} \leqslant x^{1-\alpha} \quad (0 < x \leqslant 1),$$

d'où

$$\int_{0}^{1} \frac{\sin x}{x^{\alpha}}\,dx \leqslant \int_{0}^{1} x^{1-\alpha}\,dx < \infty \quad \text{pour } \alpha < 2,$$

$$\int_{0}^{1} \frac{\sin x}{x^{\alpha}}\,dx \geqslant \frac{2}{\pi} \int_{0}^{1} x^{1-\alpha}\,dx = \infty \quad \text{pour } \alpha \geqslant 2.$$

Donc

$$\int_{0}^{1} \frac{\sin x}{x^{\alpha}}\,dx \begin{cases} \text{converge absolument pour } \alpha < 2, \\ \text{diverge pour } \alpha \geqslant 2. \end{cases} \tag{4}$$

D'autre part (cf. § 10, exemple 2)

$$\int_{1}^{\infty} \frac{\sin x}{x^{\alpha}}\,dx \begin{cases} \text{converge pour } \alpha > 0, \\ \text{converge absolument seulement pour } \alpha > 1. \end{cases} \tag{5}$$

CHAPITRE 7

# APPLICATIONS DES INTÉGRALES. MÉTHODES DES APPROXIMATIONS

## § 1. Aire en coordonnées polaires

L'aire $S$ d'une figure limitée par deux rayons vecteurs $\theta_0$ et $\theta_*$ issus du pôle $O$ et par une courbe $\Gamma$ d'équation $\rho = f(\theta)$ peut être définie de la manière suivante (fig. 80). Considérons une subdivision de l'intervalle $[\theta_0, \theta_*]$:

$$\theta_0 < \theta_1 < \ldots < \theta_n = \theta_*.$$

L'élément d'aire de la figure limitée par $\Gamma$ et par les rayons vecteurs $\theta = \theta_k$, $\theta = \theta_{k+1}$ est égal approximativement à l'aire du secteur

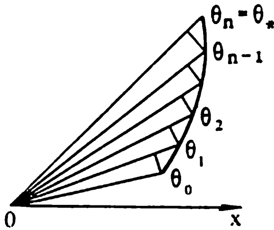


Fig. 80

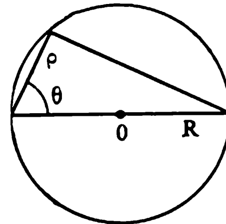


Fig. 81

limité par ces deux rayons et par le cercle de rayon $\rho_k = f(\theta_k)$, soit

$$\frac{1}{2}\rho_k^2\Delta\theta_k, \quad \Delta\theta_k = \theta_{k+1} - \theta_k.$$

Il est naturel de poser par définition

$$S = \lim_{\max \Delta\theta_k \to 0} \frac{1}{2}\sum_{k=0}^{n-1} \rho_k^2\Delta\theta_k = \frac{1}{2}\int_{\theta_0}^{\theta_*} \rho^2\, d\theta = \frac{1}{2}\int_{\theta_0}^{\theta_*} f^2(\theta)\, d\theta. \qquad (1)$$

Nous avons obtenu la formule de l'aire d'une figure en coordonnées polaires. Si la fonction $f(\theta)$ est continue, on sait que l'intégrale (1) existe et par suite est la limite de toute somme de Riemann.

Exemple. En coordonnées polaires, l'équation du cercle de la figure 81 est $\rho = 2R \cos \theta$. En vertu de (1) l'aire du disque est

$$S = 2R^2 \int_{-\pi/2}^{\pi/2} \cos^2 \theta \, d\theta = 4R^2 \int_0^{\pi/2} \frac{1+\cos 2\theta}{2} \, d\theta = \pi R^2.$$

## § 2. Volume d'un solide de révolution

Soit Γ une courbe d'équation $y = f(x)$. On suppose que $f$ est une fonction continue strictement positive. Calculons le volume $V$ d'un solide limité par la surface de révolution engendrée par Γ en tournant autour de $Ox$, et les plans $x = a$ et $x = b$.

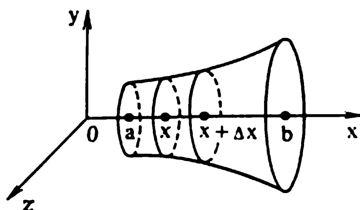



Fig. 82

Considérons la subdivision suivante de l'intervalle $[a, b]$: $a = x_0 < x_1 < \ldots < x_n = b$, et supposons que l'élément $\Delta V$ de volume limité par les plans $x = x_k$ et $x = x_{k+1}$ est approximativement égal au volume du cylindre de hauteur $\Delta x_k = x_{k+1} - x_k$ et de rayon $y_k = f(x_k)$:

$$\Delta V_k \sim \pi y_k^2 \, \Delta x_k = \pi f^2(x_k) \, \Delta x_k.$$

Le volume $V$ a pour valeur

$$V = \lim_{\max \Delta x_k \to 0} \pi \sum_{k=0}^{n-1} f^2(x_k) \, \Delta x_k = \pi \int_a^b f^2(x) \, dx. \qquad (1)$$

Nous avons ainsi établi la formule du volume d'un solide de révolution autour de $Ox$ (fig. 82).

Exemple. L'ellipsoïde de révolution (autour de l'axe $Ox$)

$$\frac{x^2}{a^2} + \frac{y^2+z^2}{b^2} \leqslant 1$$

est engendré par la courbe

$$y = b \sqrt{1 - \frac{x^2}{a^2}} \quad (-a \leqslant x \leqslant a)$$

en tournant autour de l'axe $Ox$, donc en vertu de (1) son volume a pour valeur

$$V = \pi b^2 \int_{-a}^{a} \left(1 - \frac{x^2}{a^2}\right) dx = \pi b^2 \left(x - \frac{x^3}{3a^2}\right)_{-a}^{a} = \frac{4}{3} \pi a b^2.$$

## § 3. Courbe gauche lisse. Longueur d'arc

Au § 21 du chap. 4, on a introduit la notion de courbe continue plane définie paramétriquement et notamment la notion de courbe lisse.

On se propose de compléter ces notions. Considérons une courbe plus générale dans l'espace. Les équations

$$\begin{cases} x=\varphi(t), \\ y=\psi(t), \quad a\leqslant t\leqslant b, \\ z=\chi(t), \end{cases} \tag{1}$$

où les fonctions $\varphi$, $\psi$ et $\chi$ sont continues sur $[a, b]$, définissent une *courbe gauche continue* que nous désignerons par $\Gamma$ (fig. 83). Si, de plus, les fonctions $\varphi$, $\psi$ et $\chi$ admettent des dérivées continues sur $[a, b]$ et non toutes nulles, alors on dit que $\Gamma$ est une *courbe différentiable* ou *lisse*.

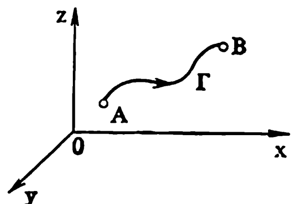



Fig. 83

On peut exprimer le fait que les dérivées $\varphi'(t)$, $\psi'(t)$ et $\chi'(t)$ ne sont pas toutes nulles quel que soit $t \in [a, b]$ au moyen de l'inégalité

$$(\varphi'(t))^2+(\psi'(t))^2+(\chi'(t))^2>0, \quad \forall t\in[a, b]. \tag{2}$$

Considérons une valeur $t = t_0$. En vertu de (2) l'un des termes $\varphi'(t_0)$, $\psi'(t_0)$ et $\chi'(t_0)$ n'est pas nul. Supposons qu'il s'agisse de $\varphi'(t_0)$. La fonction $\varphi'$ étant continue, il existe un intervalle $]t_0 - \delta, t_0 + \delta[$ sur lequel $\varphi'(t)$ est du même signe que $\varphi'(t_0)$. Donc, la fonction $x = \varphi(t)$ est strictement monotone sur cet intervalle et possède une fonction réciproque continûment dérivable $t = \varphi^{-1}(x)$, $x \in ]c, d[$, où $]c, d[$ est un voisinage du point $x_0 = \varphi(t_0)$. On obtient ainsi une petite portion $\gamma$ de courbe $\Gamma$ contenant le point $A_0 = (\varphi(t_0), \psi(t_0), \chi(t_0))$ et définie par les équations

$$y=\psi[\varphi^{-1}(x)]=\psi_1(x),$$
$$z=\chi[\varphi^{-1}(x)]=\chi_1(x),$$

où $\psi_1$ et $\chi_1$ sont des fonctions continûment dérivables de $x$ ($x, x_0 \in ]c, d[$, $x_0 = \varphi(t_0)$). Si $\psi'(t_0) \neq 0$ ou $\chi'(t_0) \neq 0$, alors par des raisonnements analogues on trouve qu'une portion $\gamma \subset \Gamma$ est définie par les équations

$$x=\varphi_1(y), \quad z=\chi_1(y)$$
$$(y, y_0\in]\lambda, \mu[, \quad y_0=\psi(t_0))$$

ou respectivement

$$x=\varphi_1(z), \quad y=\psi_1(z)$$
$$(z, z_0\in]p, q[, \quad z_0=\chi(t_0)).$$

Les équations (1) définissent non seulement la courbe $\Gamma$, mais aussi son *orientation*, c'est-à-dire le sens de croissance du paramètre $t$. La figure 83 représente une courbe lisse $\Gamma$ correspondant aux variations de $t$ sur $[a, b]$: $A = (\varphi(a), \psi(a), \chi(a))$ est l'origine de $\Gamma$, $B = (\varphi(b), \psi(b), \chi(b))$, l'extrémité, la flèche indique le sens de $\Gamma$.

Lorsque le paramètre $t$ croît continûment de $a$ à $b$, le point $(\varphi(t), \psi(t), \chi(t))$ se déplace continûment sur $\Gamma$ de l'origine $A = (\varphi(a), \psi(a), \chi(a))$ à l'extrémité $B = (\varphi(b), \psi(b), \chi(b))$. Le point courant peut revenir à une ancienne position, c'est-à-dire que pour $t_1, t_2 \in [a, b]$, $t_1 < t_2$, on aura $\varphi(t_1) = \varphi(t_2)$, $\psi(t_1) = \psi(t_2)$, $\chi(t_1) = \chi(t_2)$. On dit alors que *la courbe* $\Gamma$ *présente un point double ou de self-intersection*. La courbe $\Gamma$ est dite *fermée* si les points $A$ et $B$ sont confondus.

Considérons une fonction $t = \lambda(\tau)$, $\tau \in [c, d]$, possédant une dérivée continue non nulle sur $[c, d]$ et appliquant $[c, d]$ sur $[a, b]$. Comme $\lambda'(\tau)$ garde son signe sur $[c, d]$, deux cas seulement sont possibles:

1) $\lambda'(\tau) > 0$ et alors $\lambda(c) = a$, $\lambda(d) = b$,

2) $\lambda'(\tau) < 0$ et alors $\lambda(c) = b$, $\lambda(d) = a$.

La courbe continue $\Gamma$ peut être définie par les équations

$$\begin{cases} x = \varphi[\lambda(\tau)] = \varphi_1(\tau), \\ y = \psi[\lambda(\tau)] = \psi_1(\tau), \quad \tau \in [c, d]. \\ z = \chi[\lambda(\tau)] = \chi_1(\tau), \end{cases} \tag{1'}$$

On remarque donc qu'une courbe peut être définie paramétriquement au moyen de paramètres différents.

On remarquera aussi que les conditions (2) restent en vigueur, car, en vertu de la formule de dérivation d'une fonction composée, on a

$$(\varphi_1'(\tau))^2 + (\psi_1'(\tau))^2 + (\chi_1'(\tau))^2 = \\ = [(\varphi'(t))^2 + (\psi'(t))^2 + (\chi'(t))^2](\lambda'(\tau))^2 > 0. \tag{3}$$

Cependant, l'introduction du paramètre $\tau$ est susceptible de modifier l'orientation de $\Gamma$. Si $\lambda'(\tau) > 0$ sur $[c, d]$, la fonction $t = \lambda(\tau)$ est strictement croissante et $\lambda(c) = a$, $\lambda(d) = b$. Dans ce cas $t$ croît avec $\tau$ et passe de $\lambda(c) = a$ à $\lambda(d) = b$, c'est-à-dire que l'orientation de $\Gamma$ ne change pas: les équations (1) et (1') définissent la même courbe lisse avec la même orientation mais au moyen de paramètres différents. Si $\lambda'(\tau) < 0$ sur $[c, d]$, alors $\lambda(c) = b$ et $\lambda(d) = a$, et le paramètre $t$ décroît lorsque $\tau$ croît. Dans ce cas les équations (1') définissent la même courbe que les équations (1) mais avec une orientation contraire.

Dans les problèmes où l'on aura à tenir compte de l'orientation de la courbe, par $\Gamma$ on comprendra non seulement la courbe mais

aussi son orientation. On se rappellera aussi que les équations (1) définissent aussi bien la courbe que son orientation (le mouvement du point courant dans le sens des $t$ croissants). Si l'on remplace le paramètre $t$ par un paramètre $\tau$ ($t = \lambda(\tau)$), on obtient la courbe avec la même orientation ou l'orientation contraire selon que $\lambda'(\tau)$ est $>0$ ou $<0$. La courbe $\Gamma$ orientée dans le sens contraire est désignée par $\Gamma_-$.

Si une courbe orientée $\Gamma$ est définie par les équations (1), la courbe $\Gamma_-$ peut être définie par les équations

$$\begin{cases} x = \varphi(-\tau), \\ y = \psi(-\tau), \quad \tau \in [-b, -a]. \\ z = \chi(-\tau), \end{cases}$$

Introduisons la notion de longueur d'arc d'une courbe continue $\Gamma$. Soit donnée une courbe continue $\Gamma$ d'équations (1). Considérons une

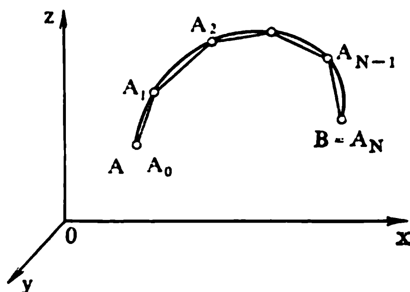



Fig. 84

subdivision de l'intervalle $[a, b]$: $a = t_0 < t_1 < t_2 < \ldots < t_N = b$. A chaque $t_k$ correspond un point $A_k \in \Gamma$ ($A_0 = A$, $A_N = B$). En reliant successivement les points $A_k$ par des segments, on obtient une ligne polygonale $\Gamma_N = A_0A_1 \ldots A_N$ inscrite dans $\Gamma$ (fig. 84). La longueur de $\Gamma_N$ est égale à la somme des longueurs $|A_kA_{k+1}|$:

$$|\Gamma_N| = \sum_{k=0}^{N-1} |A_kA_{k+1}| \tag{4}$$

*La limite (si elle existe)*

$$\lim_{\max(t_{j+1}-t_j)\to 0} |\Gamma_N| = |\Gamma| \tag{5}$$

*s'appelle longueur de l'arc* $\Gamma$. Nous l'avons désignée par $|\Gamma|$.

On démontre que pour toute courbe continue (1) la limite (5) existe. Si elle est finie, on dit que la courbe $\Gamma$ est *rectifiable*.

THÉORÈME. *Toute courbe $\Gamma$ définie par les équations* (1) *est rectifiable et sa longueur d'arc a pour valeur*

$$|\Gamma| = \int_a^b \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2 + [\chi'(t)]^2}\, dt. \tag{6}$$

Dans cet énoncé, il est important que les équations de $\Gamma$ soient définies sur un intervalle fermé $[a, b]$. Si les fonctions étaient continûment dérivables sur l'intervalle ouvert $]a, b[$ et leurs dérivées non toutes nulles, nous aurions dit aussi que les équations (1) définissent une courbe lisse, mais sans pouvoir affirmer qu'elle est rectifiable. Cependant toute portion correspondant à un intervalle fermé $[c, d] \subset ]a, b[$ est rectifiable.

DÉMONSTRATION DU THÉORÈME. Appliquons le théorème des accroissements finis aux fonctions $\varphi$, $\psi$ et $\chi$. On a ($\Delta t_k = t_{k+1} - t_k$, $\lambda_R = \max_k \Delta t_k$)

$$\Delta\varphi = \varphi(t_{k+1}) - \varphi(t_k) = \varphi'(t'_k)\,\Delta t_k,$$
$$\Delta\psi = \psi(t_{k+1}) - \psi(t_k) = \psi'(t''_k)\,\Delta t_k,$$
$$\Delta\chi = \chi(t_{k+1}) - \chi(t_k) = \chi'(t'''_k)\,\Delta t_k,$$

donc

$$|\Gamma_N| = \sum_{k=0}^{N-1} \sqrt{\Delta^2\varphi + \Delta^2\psi + \Delta^2\chi} =$$
$$= \sum_{k=0}^{N-1} \sqrt{[\varphi'(t'_k)]^2 + [\psi'(t''_k)]^2 + [\chi'(t'''_k)]^2}\,\Delta t_k =$$
$$= \sum_{k=0}^{N-1} \sqrt{[\varphi'(t'_k)]^2 + [\psi'(t'_k)]^2 + [\chi'(t'_k)]^2}\,\Delta t_k +$$
$$+ r_N \xrightarrow[\lambda_R \to 0]{} \int_a^b \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2 + [\chi'(t)]^2}\, dt \tag{7}$$

($t'_k$, $t''_k$, $t'''_k \in ]t_k, t_{k+1}[$ sont des points distincts), c'est-à-dire que la formule (6) est valable.

En effet, l'intégrant de (6) étant une fonction continue, il vient

$$\lim_{\lambda_R \to 0} \sum_{k=0}^{N-1} \sqrt{[\varphi'(t'_k)]^2 + [\psi'(t'_k)]^2 + [\chi'(t'_k)]^2}\,\Delta t_k =$$
$$= \int_a^b \sqrt{[\varphi'(t)]^2 + [\psi'(t)]^2 + [\chi'(t)]^2}\, dt.$$

D'autre part,

$$|r_N| = \left| \sum_{k=0}^{N-1} [\sqrt{[\varphi'(t_k')]^2 + [\psi'(t_k'')]^2 + [\chi'(t_k''')]^2} - \right.$$

$$\left. - \sqrt{[\varphi'(t_k')]^2 + [\psi'(t_k')]^2 + [\chi'(t_k')]^2}]\, \Delta t_k \right| \leqslant$$

$$\leqslant \sum_{k=0}^{N-1} \sqrt{[\psi'(t_k'') - \psi'(t_k')]^2 + [\chi'(t_k''') - \chi'(t_k')]^2}\, \Delta t_k \leqslant$$

$$\leqslant \frac{\varepsilon\sqrt{2}}{2(b-a)} \sum_{k=0}^{N-1} \Delta t_k < \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

Les fonctions $\psi'$ et $\chi'$ sont uniformément continues sur $[a, b]$, car continues. Donc, si $\lambda_R < \delta$, alors

$$|\psi'(t_k'') - \psi'(t_k')| < \frac{\varepsilon}{2(b-a)}, \quad |\chi'(t_k''') - \chi'(t_k')| < \frac{\varepsilon}{2(b-a)}.$$

Pour majorer $|r_N|$ on s'est servi de l'inégalité triangulaire

$$|\sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2} - \sqrt{\eta_1^2 + \eta_2^2 + \eta_3^2}| \leqslant$$

$$\leqslant \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + (\xi_3 - \eta_3)^2}.$$

Appliquons la formule (6) au calcul de la longueur d'arc de $\Gamma$ définie par les équations (1'). On a (voir (6))

$$|\Gamma_1| = \int_c^d \sqrt{(\varphi_1'(\tau))^2 + (\psi_1'(\tau))^2 + (\chi_1'(\tau))^2}\, d\tau =$$

$$= \int_c^d \sqrt{(\varphi'(\lambda(\tau)))^2 + (\psi'(\lambda(\tau)))^2 + (\chi'(\lambda(\tau)))^2}\, \lambda'(\tau)\, d\tau =$$

$$= \int_a^b \sqrt{(\varphi'(t))^2 + (\psi'(t))^2 + (\chi'(t))^2}\, dt \quad (\lambda'(\tau) > 0).$$

Donc, $|\Gamma_1| = |\Gamma|$.

Nous constatons que la formule (6) exprime intrinsèquement la longueur d'arc.

Considérons la fonction

$$s = \mu(t) = \int_a^t \sqrt{(\varphi'(u))^2 + (\psi'(u))^2 + (\chi'(u))^2}\, du \quad (a \leqslant t \leqslant b). \quad (8)$$

Elle exprime la longueur de l'arc $\overset{\frown}{AC}$, où $C$ est le point variable de l'arc $\overset{\frown}{AB} = \Gamma$, correspondant à la valeur du paramètre $t$. L'inté-

grant de (8) est une fonction continue de $u$, car la dérivée de la longueur d'arc $s$ par rapport à $t$ est égale à

$$\frac{ds}{dt}=\sqrt{(\varphi'(t))^2+(\psi'(t))^2+(\chi'(t))^2}. \tag{9}$$

Comme $\varphi'(t)$, $\psi'(t)$ et $\chi'(t)$ sont continues, $ds/dt$ l'est aussi et de plus elle est strictement positive (cf. (2)). Donc, $s = \mu(t)$ est strictement croissante sur $[a, b]$ et possède une fonction réciproque

$$t = \mu^{-1}(s), \quad 0 \leqslant s \leqslant |\Gamma|, \tag{10}$$

continûment dérivable telle que

$$\frac{dt}{ds}=[(\varphi'(t))^2+(\psi'(t))^2+(\chi'(t))^2]^{-1/2}>0.$$

Donc, la variable $s$ peut être prise pour paramètre et l'équation de $\Gamma$ peut alors s'écrire

$$\begin{cases} x=\varphi[\mu^{-1}(s)]=\varphi_*(s), \\ y=\psi[\mu^{-1}(s)]=\psi_*(s), \\ z=\chi[\mu^{-1}(s)]=\chi_*(s), \end{cases}$$

les fonctions $\varphi_*$, $\psi_*$ et $\chi_*$ étant continûment dérivables sur $[0, |\Gamma|]$.

Si l'on se place dans le plan, il faut poser $z = \chi(t) \equiv 0$ dans les raisonnements précédents. La courbe plane différentiable $\Gamma$ est définie alors par deux équations:

$$\begin{cases} x=\varphi(t) \\ y=\psi(t), \end{cases} \quad t\in[a, b],$$

où $\varphi$ et $\psi$ sont des fonctions continûment dérivables telles que

$$(\varphi'(t))^2 + (\psi'(t))^2 > 0, \quad t \in [a, b].$$

La longueur de $\Gamma$ a pour valeur

$$|\Gamma| = \int_a^b \sqrt{(\varphi'(t))^2+(\psi'(t))^2}\,dt. \tag{6'}$$

La longueur d'arc $\overset{\frown}{AC} \subset \Gamma$, où $C$ est un point de $\Gamma$, correspondant à la valeur du paramètre $t\in[a, b]$, vaut

$$s=\int_a^t \sqrt{(\varphi'(u))^2+(\psi'(u))^2}\,du, \tag{8'}$$

et

$$ds=\sqrt{(\varphi'(t))^2+(\psi'(t))^2}\,dt. \tag{9'}$$

Si $\Gamma$ est définie par une fonction continûment dérivable

$$y = f(x), \quad x \in [a, b],$$

on peut admettre que $x$ est un paramètre et l'équation de $\Gamma$ devient

$$\begin{cases} x=x, \\ y=f(x), \end{cases} \quad x\in[a, b].$$

En vertu de (6′)

$$|\Gamma| = \int_a^b \sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx.$$

La différentielle d'arc de $\Gamma$ s'exprime par la formule

$$ds = \sqrt{dx^2+dy^2}.$$

EXEMPLE 1. Trouver la longueur d'arc de la courbe $\Gamma$ : $y = \operatorname{ch} x$, $x \in [0,\ 2]$.

On a

$$|\Gamma| = \int_0^2 \sqrt{1+(\operatorname{sh} x)^2}\,dx = \int_0^2 \operatorname{ch} x\,dx = \operatorname{sh} x\,\big|_0^2 = \operatorname{sh} 2 = \frac{e^2-e^{-2}}{2}.$$

EXEMPLE 2. Trouver la longueur d'un cercle $\Gamma$ de rayon $R$. Le cercle est défini paramétriquement par les équations:

$$\begin{cases} x = R\cos t, \\ y = R\sin t, \end{cases} \quad t \in [0,\ 2\pi].$$

Alors

$$|\Gamma| = \int_0^{2\pi} \sqrt{R^2\sin^2 t + R^2\cos^2 t}\,dt = R\int_0^{2\pi} dt = 2\pi R.$$

EXEMPLE 3. Trouver la longueur d'arc de la courbe $\Gamma$: $y = \int_0^x \sqrt{2t+t^2}\,dt$, lorsque $x$ varie entre 0 et 2.

On pourrait exprimer $y$ en fonction de $x$ en calculant l'intégrale par une substitution d'Euler ou par une substitution ramenant l'intégrant à une fonction rationnelle. Or, ici on peut se dispenser d'expliciter $y$. En effet, $y' = \sqrt{2x+x^2}$. Donc,

$$|\Gamma| = \int_0^2 \sqrt{1+2x+x^2}\,dx = \int_0^2 (x+1)\,dx = \frac{(x+1)^2}{2}\Big|_0^2 = 4.$$

## § 4. Courbure et rayon de courbure d'une courbe. Développée et développante

On appelle *courbure d'un cercle* de rayon $R$ le nombre $1/R$. Ce nombre peut encore être défini comme le rapport de l'angle formé par les tangentes aux extrémités d'un arc de cercle à la longueur de cet arc. L'angle formé par les tangentes au cercle en $A$ et $B$ est égal à l'angle au centre $\alpha$ des rayons $OA$ et $OB$. La longueur $|\overset{\frown}{AB}|$ de

l'arc $\widehat{AB}$ est $R\alpha$. Donc (fig. 85)

$$\frac{\alpha}{|\widehat{AB}|}=\frac{\alpha}{R\alpha}=\frac{1}{R}.$$

Cette définition de la courbure du cercle nous suggère de définir la courbure d'une courbe lisse quelconque $\Gamma$ de la manière suivante.

Soit une courbe lisse arbitraire $\Gamma$. Comme indiqué au § 3, elle est rectifiable et l'on peut parler de la longueur de l'un quelconque de ses arcs $\widehat{AB}$. L'angle $\alpha$ $(0 \leqslant \alpha \leqslant \pi)$ formé par les tangentes à $\Gamma$ en $A$ et $B$ s'appelle *angle de contingence de l'arc* $\widehat{AB}$. Le rapport de l'angle de contingence de l'arc $\widehat{AB}$ à la longueur de ce dernier s'appelle *courbure moyenne de* $\widehat{AB}$ (fig. 86). Enfin, la *courbure de la courbe* $\Gamma$ en un de ses points $A$ est par définition la limite (finie ou infinie) du rapport de l'angle de contingence $\alpha$ de l'arc $\widehat{AB}$ à sa longueur $|\widehat{AB}| = |\Delta s|$ lorsque cette dernière tend vers 0:

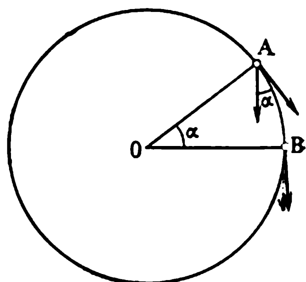


Fig. 85

$$K=\lim_{\Delta s\to 0}\frac{\alpha}{|\Delta s|}. \tag{1}$$

Donc, $K \in [0, \infty]$. Par définition, $R = 1/K$ est le *rayon de courbure* de $\Gamma$ en $A$. Dans le calcul de $R$ on convient que $0 = 1/\infty$ et $\infty = 1/0$.

Le point $O_1$ situé sur la normale à $\Gamma$ en $A$ à une distance $R = 1/K$ de $A$ dans le sens de la concavité de $\Gamma$ s'appelle *centre de courbure* de $\Gamma$ en $A$ (fig. 87 et 88). Il est évident que le centre du cercle est confondu avec le centre de courbure.

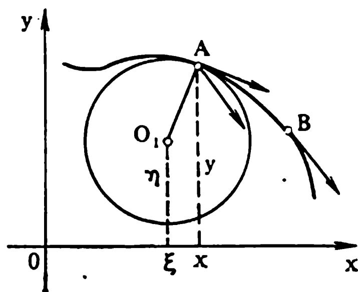


Fig. 86

On appelle *développée* de $\Gamma$ la courbe $\gamma$, lieu géométrique des centres de courbure $O_1$ de $\Gamma$. La courbe $\Gamma$ s'appelle *développante* de $\gamma$.

Soit une courbe $\Gamma$ d'équation $y = f(x)$ $(c \leqslant x \leqslant d)$. On admet que $f$ a une dérivée seconde continue. Trouvons la courbure de $\Gamma$ en $A = (x, f(x))$. Soient $\varphi_1$ et $\varphi_2$ les angles formés par les tangentes

à $\Gamma$ en $A$ et $B = (x + \Delta x,\ f(x + \Delta x))$ avec l'axe $Ox$ (fig. 86)

$$\operatorname{tg} \varphi_1 = f'(x), \quad \operatorname{tg} \varphi_2 = f'(x + \Delta x),$$
$$\alpha = |\operatorname{Arctg} f'(x) - \operatorname{Arctg} f'(x + \Delta x)|. \tag{2}$$

On a

$$\Delta s = |\widehat{AB}| = \int_x^{x+\Delta x} \sqrt{1 + (f'(u))^2}\, du. \tag{3}$$

A partir de (1), on obtient en appliquant la règle de l'Hospital

$$K = \lim_{\Delta x \to 0} \left| \frac{\operatorname{Arctg} f'(x) - \operatorname{Arctg} f'(x + \Delta x)}{\int\limits_x^{x+\Delta x} \sqrt{1 + (f'(u))^2}\, du} \right| =$$

$$= \lim_{\Delta x \to 0} \left| \frac{\dfrac{f''(x + \Delta x)}{1 + (f'(x + \Delta x))^2}}{\sqrt{1 + (f'(x + \Delta x))^2}} \right| = \frac{|f''(x)|}{(1 + (f'(x))^2)^{3/2}}.$$

D'où la formule de la courbure

$$K = \left| \frac{f''(x)}{(1 + (f'(x))^2)^{3/2}} \right|. \tag{4}$$

Si la courbe lisse $\Gamma$ est définie paramétriquement par les équations

$$\begin{cases} x = \varphi(t), \\ y = \psi(t), \end{cases} \quad a \leqslant t \leqslant b,$$

où $\varphi$ et $\psi$ sont des fonctions bicontinûment dérivables, on obtient (cf. chap. 4, § 11)

$$f'(x) = \frac{y'_t}{x'_t}, \quad f''(x) = \frac{x'_t y''_t - y'_t x''_t}{(x'_t)^3},$$

$$R = \left| \frac{(x'^2_t + y'^2_t)^{3/2}}{y'_t x''_t - x'_t y''_t} \right|, \quad K = \frac{1}{R}. \tag{5}$$

Formons l'équation paramétrique de la développée $\gamma$ d'une courbe $\Gamma$ définie par $y = f(x)$ (fig. 87, 88). On a (cf. (4))

$$\frac{1}{R} = \frac{|f''(x)|}{(1 + (f'(x))^2)^{3/2}} = \frac{f''(x) \operatorname{sgn} f''(x)}{(1 + (f'(x))^2)^{3/2}}. \tag{6}$$

Soit $O_1(\xi, \eta)$ le centre de courbure de $\Gamma$ en $A = (x, f(x))$. Ce centre est défini par le vecteur

$$\boldsymbol{\rho} = \boldsymbol{r} + R\boldsymbol{v}, \tag{7}$$

où $\boldsymbol{r}$ est le rayon vecteur de $A \in \Gamma$, $\boldsymbol{v}$, le vecteur unité normal orienté dans le sens de concavité de $\Gamma$. La courbe $\Gamma$ a pour équation vectorielle

$$\boldsymbol{r} = (x, y).$$

D'où

$$\dot{r}_x = (1,\ y'_x),\quad \ddot{r}_x = (0,\ y''_x).$$

D'autre part (cf. chap. 4, § 23, (3')),

$$\mathbf{v} = \pm\left(\frac{-y'_x}{\sqrt{1+y'^2_x}},\ \frac{1}{\sqrt{1+y'^2_x}}\right).$$

Il faut prendre le signe tel que le vecteur $\mathbf{v}$ soit orienté dans le sens de la concavité de $\Gamma$, c'est-à-dire tel que le produit scalaire $(\mathbf{v},\ \ddot{r}_x)$

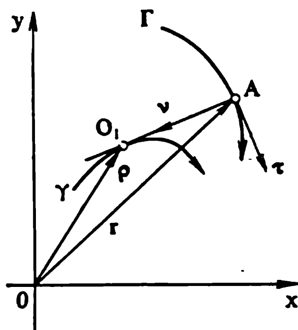



Fig. 87

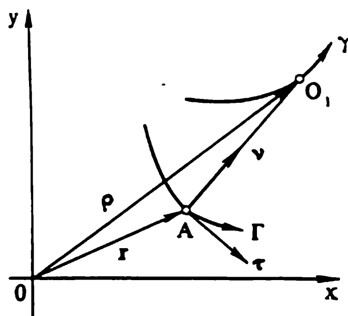



Fig. 88

soit positif:

$$(\mathbf{v},\ \ddot{r}_x) = \pm\frac{y''_x}{\sqrt{1+y'^2_x}} = y''_x\,(\operatorname{sgn} y''_x)\,(1+y'^2_x)^{-1/2}.$$

Donc

$$\mathbf{v} = \operatorname{sgn} y''_x\left(\frac{-y'_x}{\sqrt{1+y'^2_x}},\ \frac{1}{\sqrt{1+y'^2_x}}\right). \tag{8}$$

En passant aux projections dans (7) et en tenant compte de (6) et de (8), on obtient

$$\begin{aligned}\xi &= x + \frac{(1+y'^2_x)^{3/2}}{y''_x \operatorname{sgn} y''_x}\cdot\frac{-y'_x \operatorname{sgn} y''_x}{(1+y'^2_x)^{1/2}} = x - \frac{y'_x\,(1+y'^2_x)}{y''_x},\\ \eta &= y + \frac{(1+y'^2_x)^{3/2}}{y''_x \operatorname{sgn} y''_x}\cdot\frac{\operatorname{sgn} y''_x}{(1+y'^2_x)^{1/2}} = y + \frac{1+y'^2_x}{y''_x}.\end{aligned} \tag{9}$$

Montrons que *la normale à la courbe (la développante) en un point $A = (x,\ f(x))$ est tangente à la développée $\gamma$ en $O_1 = (\xi,\ \eta)$*. Il suffit pour cela de prouver que les tangentes à la courbe $\Gamma$ et à la déve-

loppée $\gamma$ aux points homologues sont orthogonales:

$$x'_x\xi'_x + y'_x\eta'_x = 1 \cdot \left[1 - y''_x \frac{1+y'^2_x}{y''_x} - y'_x \left(\frac{1+y'^2_x}{y''_x}\right)'\right] +$$
$$+ y'_x \left[y'_x + \left(\frac{1+y'^2_x}{y''_x}\right)'\right] = 0.$$

Voici une autre propriété importante de la développée. *L'accroissement du rayon de courbure de la développante est égal au signe près*

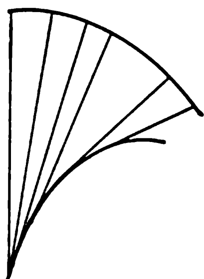


Fig. 89

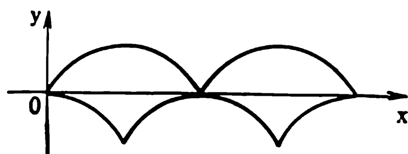



Fig. 90

*à l'accroissement de la longueur de l'arc correspondant de la développée:*

$$R_2 - R_1 = \pm \mid \sigma_2 - \sigma_1 \mid.$$

Nous glisserons sur la démonstration.

Imaginons-nous un fil enroulé sur la développante. Déroulons ce fil en le maintenant toujours tendu. Il est évident qu'il se déta-

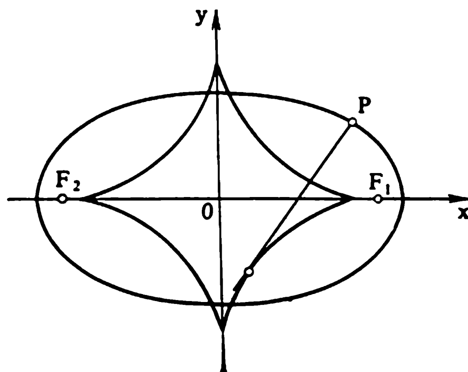



Fig. 91

chera de la développée en lui restant toujours tangent. L'extrémité libre décrira une développante (fig. 89). Comme le fil peut être de longueur arbitraire, la développée engendre une infinité de dévelop-

pantes. La longueur de fil déroulé est visiblement égale à l'accroissement du rayon de courbure de la développante. Si la courbe $\Gamma$ est définie par les équations paramétriques $x = x(t)$ et $y = y(t)$, alors la développée est donnée par les équations:

$$\xi = x - y'_t \frac{x_t'^2 + y_t'^2}{x'_t y''_t - y'_t x''_t}, \quad \eta = y + x'_t \frac{x_t'^2 + y_t'^2}{x'_t y''_t - y'_t x''_t} \tag{10}$$

(voir chap. 4, § 11).

Exemple 1. La développée de la cycloïde $x = t - \sin t$, $y = t - \cos t$ est la courbe $\xi = t + \sin t$, $\eta = -1 + \cos t$. En posant $t = \tau + \pi$, on obtient les équations

$$\xi - \pi = \tau - \sin \tau, \quad \eta + 2 = 1 - \cos \tau,$$

qui définissent la courbe de départ mais translatée (la développée de la cycloïde est une cycloïde congrue, fig. 90).

Exemple 2. La développée de l'ellipse $x = a \cos t$, $y = b \sin t$ $a \geqslant b > 0$) est une astroïde (fig. 91)

$$\xi = \frac{a^2 - b^2}{a} \cos^3 t, \quad \eta = -\frac{a^2 - b^2}{b} \sin^3 t.$$

## § 5. Aire d'une surface de révolution

Soit $\Gamma$ une courbe d'équation $y = f(x)$ dans un système de coordonnées rectangulaires. On admet que la fonction $f(x)$ possède une dérivée continue sur $[a, b]$. Calculons l'aire $S$ de la surface de révolution de $\Gamma$ autour de l'axe $Ox$. Considérons à cet effet la subdivision $a = x_0 < x_1 < \ldots < x_n = b$ de $[a, b]$ et inscrivons dans la courbe $\Gamma$ une ligne polygonale $\Gamma_n$ de sommets $(x_k, f(x_k))$. Calculons l'aire de la surface de révolution de cette dernière autour de l'axe $Ox$ (c'est-à-dire la somme des aires des surfaces latérales des cônes tronqués):

$$S_n = \pi \sum_{k=0}^{n-1} [f(x_k) + f(x_{k+1})] \sqrt{\Delta x_k^2 + \Delta y_k^2},$$

$$\Delta y_k = f(x_{k+1}) - f(x_k).$$

En appliquant le théorème des accroissements finis à $\Delta y_k$ on obtient

$$S_n = \pi \sum_{k=0}^{n-1} [f_1(x_k) + f(x_{k+1})] \sqrt{1 + (f'(\xi_k))^2}\, \Delta x_k =$$

$$= 2\pi \sum_{k=0}^{n-1} f(x_k) \sqrt{1+(f'(x_k))^2}\,\Delta x_k + R_n \to$$

$$\to 2\pi \int_a^b f(x) \sqrt{1+(f'(x))^2}\,dx$$

pour $\lambda_R = \max\limits_k \Delta x_k \to 0$; le théorème des accroissements finis dit que $\xi_k \in ]x_k, x_{k+1}[$. En effet, les fonctions $f$ et $f'$ étant continues sur $[a, b]$, la fonction $f(x)\sqrt{1 + (f'(x))^2}$ est intégrable et

$$\lim_{\lambda_R \to 0} 2\pi \sum_{k=0}^{n-1} f(x_k) \sqrt{1+(f'(x_k))^2}\,\Delta x_k = 2\pi \int_a^b f(x) \sqrt{1+(f'(x))^2}\,dx.$$

D'autre part

$$|R_n| = \Big|\pi \sum_{k=0}^{n-1} \{[f(x_k)+f(x_{k+1})] \sqrt{1+(f'(\xi_k))^2} -$$

$$-2f(x_k)\sqrt{1+(f'(x_k))^2}\}\,\Delta x_k\Big| =$$

$$= \Big|\pi \sum_{k=0}^{n-1} \{(f(x_k)+f(x_{k+1}))\,(\sqrt{1+(f'(\xi_k))^2} - \sqrt{1+(f'(x_k))^2}) +$$

$$+(f(x_{k+1})-f(x_k))\sqrt{1+(f'(x_k))^2}\}\,\Delta x_k\Big| \leqslant$$

$$\leqslant \pi \sum_{k=0}^{n-1} \{(|f(x_k)|+|f(x_{k+1})|)\,|f'(\xi_k)-f'(x_k)| +$$

$$+\sqrt{1+(f'(x_k))^2}\,|f(x_{k+1})-f(x_k)|\}\,\Delta x_k.$$

Etant continues sur $[a, b]$, les fonctions $f$ et $f'$ sont bornées et uniformément continues sur $[a, b]$. Donc $|f| \leqslant M$, $\sqrt{1 + (f'(x))^2} \leqslant M$ et pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$|f(x_{k+1}) - f(x_k)| < \frac{\varepsilon}{3\pi M(b-a)},$$

$$|f'(x_k) - f'(\xi_k)| \leqslant \frac{\varepsilon}{3\pi M(b-a)},$$

pour $\lambda_R < \delta$. D'où

$$|R_n| \leqslant \pi \sum_{k=0}^{n-1} \left\{2M\frac{\varepsilon}{3\pi M(b-a)} + M\frac{\varepsilon}{3\pi M(b-a)}\right\}\Delta x_k =$$

$$= \frac{\varepsilon}{b-a}\sum_{k=0}^{n-1}\Delta x_k = \frac{\varepsilon}{b-a}(b-a) = \varepsilon,$$

c'est-à-dire que $R_n \to 0$ avec $\lambda_R$. Donc, l'aire de la surface de révolution est égale à

$$S = 2\pi \int_a^b f(x) \sqrt{1 + (f'(x))^2}\, dx. \qquad (1)$$

EXEMPLE. Trouver l'aire $S$ de la surface de révolution de l'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a > b)$$

autour de l'axe $Ox$ (l'aire de la surface de l'ellipsoïde de révolution).

SOLUTION. L'équation de la moitié supérieure de l'ellipse est

$$y = \frac{b}{a} \sqrt{a^2 - x^2} \quad (|x| \leqslant a),$$

d'où

$$y' = -\frac{b}{a} \frac{x}{\sqrt{a^2 - x^2}}.$$

Donc

$$S = 2\pi \int_{-a}^{a} \frac{b}{a} \sqrt{a^2 - x^2} \sqrt{1 + \left(\frac{b}{a}\right)^2 \frac{x^2}{a^2 - x^2}}\, dx =$$

$$= \frac{4\pi b}{a^2} \int_0^a \sqrt{a^4 - (a^2 - b^2)\, x^2}\, dx = (u = x\sqrt{a^2 - b^2}) =$$

$$= -\frac{4\pi b}{a^2 \sqrt{a^2 - b^2}} \int_0^{a\sqrt{a^2 - b^2}} \sqrt{a^4 - u^2}\, du =$$

$$= \frac{4\pi b}{a^2 \sqrt{a^2 - b^2}} \left\{ \frac{u}{2} \sqrt{a^4 - u^2} + \frac{a^4}{2} \operatorname{Arcsin} \frac{u}{a^2} \right\}_0^{a\sqrt{a^2 - b^2}} =$$

$$= 2\pi b^2 + \frac{2\pi b a^2}{\sqrt{a^2 - b^2}} \operatorname{Arcsin} \frac{\sqrt{a^2 - b^2}}{a}.$$

Lorsque $b = a$, on obtient une sphère de rayon $a$ dont l'aire est $S = 4\pi a^2$.

## § 6. Formule d'interpolation de Lagrange

Posons le problème suivant. On demande un polynôme $L_n(x)$ de degré $\leqslant n$ qui prenne les mêmes valeurs qu'une fonction $f(x)$ en des points $x_0, x_1, \ldots, x_n$ donnés, c'est-à-dire que

$$f(x_k) = L_n(x_k) \quad (k = 0, 1, \ldots, n).$$

Le polynôme $L_n(x)$ est unique. Si l'on suppose qu'il existe un autre polynôme $\overline{L}_n(x)$ jouissant des mêmes propriétés, alors la différence

$L_n(x) - \overline{L}_n(x)$ s'annule aux points $x_0, \ldots, x_n$. Etant de degré $\leqslant n$ et possédant $n+1$ zéros, le polynôme $L_n(x) - \overline{L_n}(x)$ est identiquement nul, et par suite, $L_n(x) \equiv \overline{L}_n(x)$.

De l'unicité il résulte que si la fonction $f(x)$ est un polynôme de degré $n$, alors elle est confondue avec $L_n(x)$ pour tous les $x$ ($f(x) \equiv \equiv L_n(x)$).

Trouvons un polynôme $Q_{n,k}(x)$ de degré $n$, nul en $x_i \neq x_k$ et égal à 1 au point $x_k$. On a de toute évidence

$$Q_{n,k}(x) = A(x - x_0) \ldots (x - x_{k-1})(x - x_{k+1}) \ldots (x - x_n),$$

où la constante $A$ se détermine à partir de la condition

$$1 = Q_{n,k}(x_k) = A \prod_{\substack{i=0 \\ i\neq k}}^{n} (x_k - x_i), \quad \text{i.e.} \quad A = \prod_{\substack{i=0 \\ i\neq k}}^{n} (x_k - x_i)^{-1}.$$

Le polynôme cherché est donc de la forme

$$Q_{n,k}(x) = \prod_{\substack{i=0 \\ i\neq k}}^{n} \frac{(x - x_i)}{(x_k - x_i)}.$$

Si l'on introduit le symbole de Kronecker

$$\delta_{ki} = \begin{cases} 0 & \text{si } k \neq i, \\ 1 & \text{si } k = i, \end{cases}$$

alors

$$Q_{n,k}(x_i) = \delta_{ki}.$$

La solution du problème posé est donnée par le polynôme

$$L_n(x) = \sum_{k=0}^{n} Q_{n,k}(x) f(x_k), \tag{1}$$

car

$$L_n(x_i) = \sum_{k=0}^{n} Q_{n,k}(x_i) f(x_k) = \sum_{k=0}^{n} \delta_{ki} f(x_k) = f(x_i)$$
$$(i = 0, 1, \ldots, n).$$

Le polynôme (1) s'appelle *polynôme d'interpolation de Lagrange*.

Comme pour le reste de la formule de Taylor, on montre que si $f(x)$ possède une dérivée d'ordre $n+1$, alors

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \tag{2}$$

où

$$\omega_{n+1}(x) = \prod_{i=0}^{n} (x - x_i)$$

et $\xi$ est un point appartenant au plus petit intervalle contenant les points $x_0, x_1, \ldots, x_n, x$. En effet, posons

$$f(x) - L_n(x) = K\omega_{n+1}(x), \qquad (3)$$

où $K$ dépend de $x$. Soit

$$\varphi(z) = f(z) - L_n(z) - K\omega_{n+1}(z),$$

où $K$ a la même valeur que dans (3). Il est évident que $\varphi(x) = 0$, $\varphi(x_i) = 0$ ($i = 0, 1, \ldots, n$). Supposons par exemple que $x_0 < x < x_1 < \ldots < x_n$. En appliquant le théorème de Rolle à la fonction $\varphi$ sur les intervalles $[x_0, x]$, $[x, x_1]$, $\ldots$, $[x_{n-1}, x_n]$, on trouve que la dérivée $\varphi'(x)$ s'annule à l'intérieur de chacun d'eux. En appliquant ensuite le théorème de Rolle successivement aux fonctions $\varphi', \ldots, \varphi^{(n)}$, on trouve qu'il existe un point $\xi$, appartenant au plus petit intervalle contenant les points $x, x_0, x_1, \ldots, x_n$, en lequel $\varphi^{(n+1)}(\xi) = 0$. Or,

$$\varphi^{(n+1)}(z) = f^{(n+1)}(z) - K(n+1)!\ .$$

En faisant $z = \xi$, on obtient $K = \frac{f^{(n+1)}(\xi)}{(n+1)!}$. Donc

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_{n+1}(x),$$

ceci prouve (2).

Le polynôme d'interpolation de Lagrange est utilisé dans le calcul approché des dérivées d'une fonction $f(x)$ dont on connaît les valeurs en $x_0, x_1, \ldots, x_n$. Plus exactement, on admet que

$$f^{(k)}(x) \approx L_n^{(k)}(x).$$

Si par exemple l'on connaît les valeurs de $f(x)$ aux points $x_0$ et $x_1$ et que l'on forme le polynôme de Lagrange $L_1(x)$ correspondant, on trouve que

$$f'(x) \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Dans les paragraphes suivants, on indiquera comment appliquer le polynôme de Lagrange au calcul approché d'une intégrale définie.

## § 7. Calcul approché de l'intégrale par la méthode des rectangles et des trapèzes

Soit à calculer l'intégrale définie d'une fonction $f$ continue sur un intervalle $[a, b]$. Si l'on connaît une primitive de $f$, on se sert naturellement de la formule de Newton-Leibniz. Sinon, et c'est souvent le cas, on fait appel au calcul approché.

Le procédé le plus simple de calcul approché d'une intégrale définie découle de la définition même de cette intégrale. Divisons

l'intervalle $[a, b]$ en parties égales par les points

$$x_k = a + k\frac{b-a}{N} \quad (k = 0, 1, \ldots, N) \tag{1}$$

et posons

$$\int_a^b f(x)\,dx \approx \frac{b-a}{N} \sum_{k=0}^{N-1} f\left(\frac{x_k + x_{k+1}}{2}\right), \tag{2}$$

où le symbole $\approx$ représente l'égalité approchée.

L'expression (2) s'appelle *formule d'intégration par la méthode des rectangles.* Dans le cas de la figure 92, l'aire cherchée de la figure

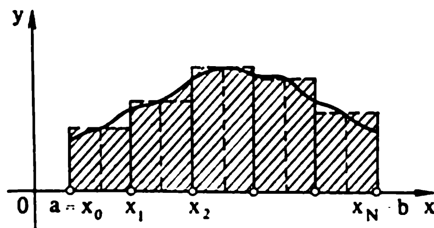



Fig. 92

limitée par la courbe $y = f(x)$, l'axe $Ox$ et les droites $x = a$ et $x = b$ est approximativement égale à la somme des aires des rectangles.

Nous savons que, si la fonction $f(x)$ est continue sur $[a, b]$, la limite pour $N \to \infty$ du second membre de l'égalité approchée (2) est égale au premier, ce qui nous permet de considérer que pour $N$ grand l'erreur de la formule d'intégration (2), c'est-à-dire la valeur absolue de la différence de ses deux membres, est petite. Le problème de l'estimation de cette erreur se pose tout naturellement. Nous verrons plus loin comment établir cette estimation si la fonction $f$ est astreinte, outre à être continue, à vérifier certaines conditions de dérivabilité (c'est-à-dire à posséder un certain nombre de dérivées).

Il importe de remarquer que si la fonction $f(x)$ est linéaire, i.e. $f(x) = Ax + B$, alors *la formule* (2) *est exacte.* Comme une fonction linéaire est un polynôme du premier degré, on peut dire que *la formule d'intégration* (2) *est exacte pour les polynômes de degré* $\leqslant 1$.

Citons une autre méthode de calcul approché d'une intégrale définie, appelée *méthode des trapèzes.* Elle consiste à diviser l'intervalle $[a, b]$ en parties égales par les points (1) et admettre approximativement que

$$\int_a^b f(x)\,dx \approx \frac{b-a}{N}\left(\frac{f(x_0)+f(x_1)}{2} + \right.$$

$$+\frac{f(x_1)+f(x_2)}{2}+\ldots+\frac{f(x_{N-1})+f(x_N)}{2}\bigg)=$$

$$=\frac{b-a}{2N}[f(x_0)+2f(x_1)+2f(x_2)+\ldots+2f(x_{N-1})+f(x_N)]. \quad (3)$$

Dans la formule des trapèzes, la surface de la figure mixtiligne est approximativement recouverte par celles des trapèzes (fig. 93).

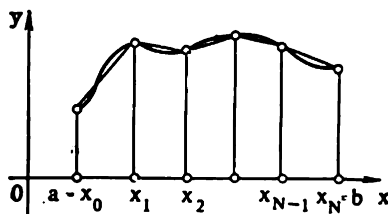



Fig. 93

Il est important de signaler que *la formule des trapèzes est exacte pour les fonctions linéaires* $Ax + B$ ($A$ et $B$ sont des constantes), c'est-à-dire pour les polynômes de degré $\leqslant 1$ ; si l'on substitue une telle fonction à $f(x)$ dans (3), on obtient une égalité. De ce point de vue la formule des trapèzes ne présente aucun avantage sur celle des rectangles: elles sont toutes deux exactes pour les fonctions linéaires.

Désignons par $R_N(f)$ la différence des deux membres de la formule d'intégration et appelons-la *reste de la formule d'intégration.*

Si la fonction $f$ admet une dérivée $f'$ continue par morceaux et telle que $|f'(x)| \leqslant M_1$, alors le reste de la formule des rectangles (2) vérifie l'inégalité

$$|R_N(f)| \leqslant \frac{M_1(b-a)^2}{4N}, \quad (2')$$

et le reste de la formule des trapèzes, l'inégalité

$$|R_N(f)| \leqslant \frac{M_1(b-a)^2}{4N}. \quad (3')$$

A noter que les constantes sont calculées ici exactement, elles ne peuvent être diminuées. L'estimation (2') est établie plus bas. Les autres sont données sans démonstration.

Nous remarquons que dans les deux cas, $R_N(f) = O(N^{-1})$ pour les fonctions admettant une dérivée bornée (cf. chap. 3, § 10, (14)).

Pour les fonctions $f(x)$ telles que $|f''(x)| \leqslant M_2$ sur $[a, b]$ on a la majoration

$$|R_N(f)| \leqslant \frac{M_2(b-a)^3}{12N^2}$$

qui est vraie pour les formules des rectangles et des trapèzes. Maintenant $R_N(f) = O(N^{-2})$ pour les deux formules.

Il s'avère que $R_N(f) = O(N^{-2})$ pour les fonctions $f$ telles que $|f^{(l)}(x)| \leqslant \leqslant M$ $(l > 2)$, c'est-à-dire que la précision ne s'améliore pas.

L'explication de ce phénomène réside dans le fait que les deux formules d'intégration ne sont exactes que pour les fonctions linéaires.

Si une fonction $f$ est telle que $f'''(x) \leqslant M$, on peut établir une formule d'intégration dont le reste $R_N(f) = O(N^{-3})$. Cette formule doit être exacte pour les polynômes du second degré. Or, si elle n'est pas exacte pour les polynômes du troisième degré, alors le reste $R_N(f) = O(N^{-3})$ pour les fonctions possédant une dérivée quatrième bornée.

Le phénomène décrit sera illustré sur l'exemple de la formule de Simpson [1]) au § 8.

Prouvons l'estimation (2'). Posons $h = \frac{b-a}{N}$, $\xi_k = \frac{x_k + x_{k+1}}{2}$, $x_k$ sont les points de (1). Alors

$$|R_N(f)| = \left| \int_a^b f(x)\,dx - h \sum_0^{N-1} f(\xi_k) \right| = \left| \sum_0^{N-1} \int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} f(x)\,dx - h \sum_0^{N-1} f(\xi_k) \right| \leqslant$$

$$\leqslant \sum_0^{N-1} \left| \int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} f(x)\,dx - hf(\xi_k) \right| = \sum_0^{N-1} \left| \int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} [f(x) - f(\xi_k)]\,dx \right|,$$

puisque

$$\int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} f(\xi_k)\,dx = hf(\xi_k).$$

En appliquant le théorème des accroissements finis à l'intégrant et compte tenu de ce que $|f'(x)| \leqslant M_1$, on obtient

$$|R_N(f)| \leqslant \sum_0^{N-1} \left| \int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} f'(\theta_k)(x - \xi_k)\,dx \right| \leqslant$$

$$\leqslant \sum_0^{N-1} \int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} |f'(\theta_k)|\,|x - \xi_k|\,dx \leqslant M_1 \sum_0^{N-1} \int_{\xi_k - \frac{h}{2}}^{\xi_k + \frac{h}{2}} |x - \xi_k|\,dx,$$

où $\theta_k$ est compris entre $x$ et $\xi_k$. La substitution $x - \xi_k = t$ nous donne

$$|R_N(f)| \leqslant M_1 \sum_0^{N-1} \int_{-h/2}^{h/2} |t|\,dt = 2M_1 \cdot N \int_0^{h/2} t\,dt =$$

$$= 2M_1 N \frac{t^2}{2} \Big|_0^{h/2} = \frac{M_1 N h^2}{4} = \frac{M_1 (b-a)^2}{4N}.$$

---

[1]) Thomas Simpson, mathématicien anglais (1710-1761).

## § 8. Formule de Simpson

Soit à calculer approximativement l'intégrale de la fonction continue $f(x)$:

$$\int_a^b f(x)\,dx. \quad (1)$$

On cherchera la valeur approchée de l'intégrale sous forme de la somme

$$\int_a^b f(x)\,dx \approx \sum_{k=0}^{n} p_k f(x_k), \quad (2)$$

où $p_0, p_1, \ldots, p_n$ et $x_0, x_1, \ldots, x_n \in [a, b]$ sont des nombres donnés.

La formule (2) s'appelle formule d'intégration de *nœuds* $x_k$ et de *coefficients de pondération* $p_k$.

Quand on établit des formules approchées, on exige que la formule (2) soit exacte pour les polynômes de degré $n$. Cette condition sera remplie si pour valeur approchée de l'intégrale (1) on prend l'intégrale définie du polynôme d'interpolation de Lagrange de la fonction $f$:

$$\int_a^b f(x)\,dx \approx \int_a^b L_n(x)dx = \int_a^b \sum_{k=0}^{n} Q_{n,k}(x) f(x_k)\,dx = \sum_{k=0}^{n} p_k f(x_k), \quad (3)$$

$$p_k = \int_a^b Q_{n,k}(x)\,dx \quad (k=0, 1, \ldots, n), \quad Q_{n,k}(x) = \prod_{\substack{i=0 \\ i\neq k}}^{n} \frac{(x-x_i)}{(x_k-x_i)},$$

car si $f(x)$ est un polynôme de degré $n$, alors $f(x) \equiv L_n(x)$.

Etablissons la formule (3) pour $n=2$ et pour les nœuds $x_0=a$, $x_1=\frac{a+b}{2}$, $x_2=b$. On a

$$Q_{2,0}(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} = \frac{(x-b)(2x-a-b)}{(b-a)^2} = \frac{2(x-b)^2}{(b-a)^2} + \frac{x-b}{b-a},$$

$$Q_{2,1}(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} = \frac{-4(x-a)(x-b)}{(b-a)^2} =$$

$$= -4\frac{(x-b)^2}{(b-a)^2} - 4\frac{x-b}{b-a}.$$

$$Q_{2,2}(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} = \frac{(x-a)(2x-a-b)}{(b-a)^2} = \frac{2(x-a)^2}{(b-a)^2} - \frac{x-a}{b-a}.$$

Donc

$$p_0 = \int_a^b Q_{2,0}(x)\,dx = \left[\frac{2(x-b)^3}{3(b-a)^2} + \frac{(x-b)^2}{2(b-a)}\right]_a^b = \frac{b-a}{6}.$$

Par des raisonnements analogues, on trouve

$$p_1 = \int_a^b Q_{2,1}(x)\,dx = \frac{2(b-a)}{3}, \quad p_2 = \int_a^b Q_{2,2}(x)\,dx = \frac{b-a}{6}.$$

Et la formule (3) pour $n=2$ s'écrit

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right]. \tag{4}$$

C'est la formule simple de Simpson correspondant à l'intervalle $[a, b]$.

Géométriquement, la formule (4) signifie que nous avons remplacé l'aire de la figure mixtiligne définie par la fonction $f(x)$ sur

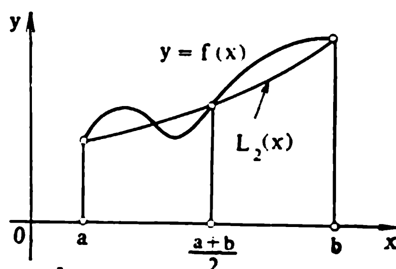



Fig. 94

$[a, b]$ par l'aire de la figure située sous le graphe de la parabole (fig. 94):

$$y = L_2(x) = f(a)\,Q_{2,0}(x) + f\left(\frac{a+b}{2}\right)Q_{2,1}(x) + f(b)\,Q_{2,2}(x).$$

On rappelle que la formule (4) est, par construction, exacte pour les polynômes du second degré. Il se trouve cependant qu'elle est aussi exacte pour les polynômes du troisième degré. En effet, $\int_a^b x^3\,dx = \frac{b^4-a^4}{4}$. On obtient le même résultat à l'aide de la

formule (4):

$$\frac{b-a}{6}\left[a^3+4\left(\frac{a+b}{2}\right)^3+b^3\right]=$$

$$=\frac{b-a}{6}\left[(a+b)(a^2-ab+b^2)+\frac{(a+b)^3}{2}\right]=$$

$$=\frac{b^2-a^2}{6}\left[\frac{3(a^2+b^2)}{2}\right]=\frac{b^4-a^4}{4}.$$

Donc, *la formule* (4) *est exacte pour les polynômes de degré* $\leqslant 3$.

Si l'on divise l'intervalle $[a, b]$ en $2N$ parties égales par les points

$$x_k=a+\frac{b-a}{2N}k \quad (k=0,\ 1,\ \ldots,\ 2N)$$

et qu'on applique la formule (4) aux intervalles $[x_0,\ x_2]$, $[x_2, x_4]$, ... on obtient la *formule* (*complexe*) *de Simpson*

$$\int_a^b f(x)\,dx \approx \frac{b-a}{6N}\{f(x_0)+4f(x_1)+2f(x_2)+$$

$$+4f(x_3)+\ldots+2f(x_{2N-2})+4f(x_{2N-1})+f(x_{2N})\}. \quad (5)$$

Que l'on se serve de la formule de Simpson ou de celle des rectangles, les calculs présentent le même degré de complexité. Mais si la fonction $f(x)$ est suffisamment lisse, il y a avantage à se servir de la formule de Simpson, car elle est plus précise.

Si la fonction $f(x)$ possède sur $[a, b]$ une dérivée seconde continue bornée:

$$|f''(x)| \leqslant M_2$$

et ne possède pas de dérivée troisième ou en possède une que l'on ne peut majorer, alors il est recommandé d'utiliser la formule des trapèzes, ou mieux encore la formule de Simpson, pour calculer l'intégrale $\int_a^b f(x)\,dx$.

On démontre que l'erreur de la formule des trapèzes (§ 7, (3)) est:

$$\frac{1}{12}\cdot\frac{(b-a)^3}{N^2}M_2 \quad (6)$$

et celle de la formule de Simpson (5)

$$\frac{1}{81}\cdot\frac{(b-a)^3}{N^2}M_2.$$

Si la fonction $f(x)$ possède sur $[a, b]$ une dérivée quatrième continue et bornée

$$|f^{(4)}(x)| \leqslant M_4,$$

il est alors conseillé d'appliquer la formule de Simpson. L'erreur sera alors:

$$\frac{1}{2880}\cdot\frac{(b-a)^5}{N^4}M_4. \quad (7)$$

Si l on s'était servi de la formule des trapèzes, l'erreur aurait été comme toujours de l'ordre de $N^{-2}$, c'est-à-dire moins bonne que (7).

EXEMPLE. Calculer l'intégrale

$$I=\int_0^1 \sqrt{1+x^4}\,dx.$$

Cette intégrale ne s'exprime pas par des fonctions élémentaires.

Calculons-la approximativement. Partageons l'intervalle [0, 1] en dix parties égales et essayons les diverses formules.

Désignons les points de subdivision de [0, 1] par $x_0 = 0$, $x_1 = 0{,}1$, . . ., $x_9 = 0{,}9$, $x_{10} = 1$. Calculons les valeurs approchées de la fonction $f(x) = \sqrt{1+x^4}$ en ces points:

$$f(0) = 1,\ f(x_1) = 1{,}00005,\ f(x_2) = 1{,}00080,$$
$$f(x_3) = 1{,}00404,\ f(x_4) = 1{,}01272,\ f(x_5) = 1{,}03078,$$
$$f(x_6) = 1{,}06283,\ f(x_7) = 1{,}11360,\ f(x_8) = 1{,}18727,$$
$$f(x_9) = 1{,}28690,\ f(x_{10}) = 1{,}41421.$$

La formule des trapèzes nous donne:

$$I \approx \frac{1}{20}[f(x_0)+2f(x_1)+\ \dots\ +2f(x_9)+f(x_{10})]=\frac{21{,}81219}{20}=1{,}09061.$$

La fonction $f(x) = \sqrt{1+x^4}$ possède autant de dérivées continues sur l'intervalle [0, 1] qu'on le veut. On a indiqué plus haut que l'existence des dérivées d'ordre $>2$ ne modifiait pas la précision de la formule des trapèzes. Donc, on calculera l'erreur de cette formule en utilisant la dérivée seconde continue

$$f''(x) = 2x^2(3+x^4)/(1+x^4)^{3/2}.$$

Comme $M_2 = \max\limits_x f''(x) = 2\sqrt{2}$, le reste de la formule des trapèzes est

$$R_{10}(f) \leqslant \frac{M_2}{12N^2}=\frac{2\sqrt{2}}{12\cdot 10^2}\approx 0{,}002357.$$

Donc,

$$I = 1{,}0906 \pm 0{,}0024.$$

La formule de Simpson nous donne ($2N = 10$)

$$I \approx \frac{1}{30}[f(x_0)+4f(x_1)+2f(x_2)+4f(x_3)+2f(x_4)+4f(x_5)+$$
$$+2f(x_6)+4f(x_7)+2f(x_8)+4f(x_9)+f(x_{10})]=\frac{32{,}68473}{30}=1{,}08949.$$

Calculons le reste de la formule de Simpson en utilisant le fait que $f(x)$ possède une dérivée quatrième continue (l'existence des dérivées d'ordre supérieur à 4 ne modifie pas la précision de la formule)

$$f^{(4)}(x) = 12(1-14x^4+5x^8)/(1+x^4)^{7/2}.$$

Comme $M_4 = \max\limits_x |f^{(4)}(x)| \leqslant 15\sqrt{2}$, il vient

$$R_5(f) \leqslant \frac{M_4}{2880N^4}\leqslant\frac{15\sqrt{2}}{2880\cdot 5^4}\leqslant 0{,}000012 < 0{,}00002.$$

Donc,

$$I = 1{,}08949 \pm 0{,}00002,$$

c'est-à-dire que la formule de Simpson est bien plus précise que la formule des trapèzes pour les fonctions suffisamment lisses et pour $N$ grand.

CHAPITRE 8

# CALCUL DIFFÉRENTIEL POUR FONCTIONS DE PLUSIEURS VARIABLES

## § 1. Notions préliminaires

La notion de fonction de plusieurs variables a été introduite à la fin du § 1 du chap. 3. Pour étudier les fonctions $z = f(x) = f(x_1, \ldots, x_n)$ de $n$ variables, c'est-à-dire les fonctions de points $x = (x_1, \ldots, x_n)$ d'un espace $R_n$ à $n$ dimensions, le lecteur doit connaître les notions élémentaires de la théorie d'un espace à $n$ dimensions [1]).

La distance [1]) de deux points $x = (x_1, \ldots, x_n)$ et $x' = (x'_1, \ldots, x'_n)$ de $R_n$ est donnée par la formule

$$|x - x'| = \sqrt{\sum_{j=1}^{n} (x_j - x'_j)^2}.$$

Signalons que si $x$, $y$, $z$ sont trois points de l'espace $R_n$, ils vérifient les inégalités

$$|x - y| \leqslant |x - z| + |y - z|, \tag{1}$$

$$||x| - |y|| \leqslant |x - y|. \tag{2}$$

Ces inégalités sont dites *inégalités triangulaires*. Elles admettent une interprétation géométrique dans un espace à trois dimensions. Les points $x$, $y$ et $z$ sont les sommets d'un triangle dont les côtés ont pour longueurs $|x - y|$, $|x - z|$, $|y - z|$. On sait du cours de géométrie que la longueur d'un côté d'un triangle est comprise entre la différence et la somme des deux autres côtés. C'est ce que traduisent les inégalités (1) et (2).

Par rapport aux coordonnées des points $x$, $y$ et $z$, l'inégalité (1) s'écrit:

$$\left(\sum_{j=1}^{n} (x_j - y_j)^2\right)^{1/2} \leqslant \left(\sum_{j=1}^{n} (x_j - z_j)^2\right)^{1/2} + \left(\sum_{j=1}^{n} (z_j - y_j)^2\right)^{1/2}. \tag{1'}$$

Ceci est l'*inégalité de Minkowski* [1]).

L'inégalité (2) découle aussitôt de (1). En effet,

$$|x| = |x - y + y| \leqslant |x - y| + |y|,$$
$$|y| = |y - x + x| \leqslant |x - y| + |x|,$$

[1]) Voir § 6 du chap. 10.

donc

$$-\mid x-y\mid \leqslant \mid x\mid - \mid y\mid \leqslant \mid x-y\mid,$$

c'est-à-dire qu'on a (2).

Par définition, une suite

$$x^1 = (x_1^1, \ldots, x_n^1),$$
$$x^2 = (x_1^2, \ldots, x_n^2),$$
$$x^3 = (x_1^3, \ldots, x_n^3),$$
$$\cdots\cdots\cdots\cdots$$

de points de $R_n$ *converge vers un point* $x^0 \in R_n$ si

$$\mid x^k - x^0 \mid \to 0 \quad (k \to \infty), \tag{3}$$

c'est-à-dire si

$$\sqrt{\sum_{j=1}^{n} (x_j^k - x_j^0)^2} \to 0. \tag{3'}$$

On écrit alors

$$\lim x^k = \lim_{k\to\infty} x^k = x^0 \quad \text{ou} \quad x^k \to x^0$$

et on dit que le point $x^0$ est la limite de la suite de points $x^k$. De (3′) il résulte que

$$x_j^k \to x_j^0, \quad k \to \infty \quad (j = 1, \ldots, n) \tag{3''}$$

et réciproquement. Donc, dire qu'une suite de points $x^k$ converge vers un point $x^0$ revient à dire que chaque coordonnée $x_j^k$ du point variable $x^k$ converge vers la coordonnée correspondante $x_j^0$ de $x^0$.

On a

$$\lim (x^k \pm y^k) = \lim x^k \pm \lim y^k,$$
$$\lim (\alpha x^k) = \alpha \lim x^k,$$
$$\lim (\alpha_k x) = x \lim \alpha_k \quad (k \to \infty),$$

où $\alpha$ et $\alpha_k$ sont des nombres. En effet, si $\lim x^k = x^0$, $\lim y^k = y^0$, $\lim \alpha_k = \alpha_0$, alors

$$\mid x^k \pm y^k - (x^0 \pm y^0) \mid \leqslant \mid x^k - x^0 \mid + \mid y^k - y^0 \mid \to 0,$$
$$\mid \alpha x^k - \alpha x^0 \mid = \mid \alpha \mid \mid x^k - x^0 \mid \to 0,$$
$$\mid \alpha_k x - \alpha_0 x \mid = \mid \alpha_k - \alpha_0 \mid \mid x \mid \to 0.$$

Une *courbe continue* $\Gamma$ de $R_n$ est un ensemble de points

$$x = x(t)$$

dépendant d'un paramètre $t \in [a, b]$ et tels que

$$\mid x(t) - x(t_0) \mid \underset{t\to t_0}{\to} 0, \quad t, t_0 \in [a, b]. \tag{4}$$

Par rapport aux coordonnées du point $x = x(t)$ de $\Gamma$, cette relation s'écrit:

$$x_j = x_j(t), \quad t \in [a, b] \quad (j = 1, \ldots, n),$$

où $x_j(t)$ sont des fonctions continues de $t$. En effet,

$$|x(t)-x(t_0)| = \left(\sum_{j=1}^{n} [x_j(t)-x_j(t_0)]^2\right)^{1/2}.$$

Si le premier membre de cette égalité tend vers 0 pour $t \to t_0$, il est évident que pour tout $j = 1, \ldots, n$

$$x_j(t) \to x_j(t_0), \quad t \to t_0, \tag{5}$$

et réciproquement, si (5) est réalisée pour tout $j = 1, \ldots, n$, alors on a (4).

## § 2. Ensemble ouvert

Soit un point $x^0 = (x_1^0, \ldots, x_n^0)$ d'un espace $R_n$ à $n$ dimensions. On appelle *boule fermée* de rayon $r > 0$ et de centre $x^0$ l'ensemble des po nts $x = (x_1, \ldots, x_n) \in R_n$ tels que

$$|x-x^0| = \left[\sum_1^n (x_j-x_j^0)^2\right]^{1/2} \leqslant r.$$

On appelle *boule ouverte* de rayon $r$ et de centre $x^0$ l'ensemble des points $x$ tels que

$$|x-x^0| < r.$$

Dans $R_n$, on appelle *parallélotope fermé* l'ensemble des points $x$ dont les coordonnées vérifient les inégalités $a_j \leqslant x_j \leqslant b_j$ ($j = 1, \ldots, n$). Si $n = 3$, on obtient un parallélépipède rectangle dont les faces sont parallèles aux axes de coordonnées rectangulaires ($Ox_1$, $Ox_2$, $Ox_3$). On appelle *parallélotope ouvert* l'ensemble des points $x = (x_1, \ldots, x_n)$ de $R_n$ vérifiant les inégalités strictes $a_j < x_j < b_j$ ($j = 1, \ldots, n$).

L'ensemble des points $x$ dont les coordonnées vérifient les inégalités $|x_j - x_j^0| \leqslant a$ ($j = 1, \ldots, n$), où $a > 0$ est un nombre donné, est un *cube fermé* de centre $x^0$ et de côté $2a$. Il est évident que pour $n = 3$, c'est un cube dont les faces sont parallèles aux axes de coordonnées (rectangulaires). Enfin, on appelle *cube ouvert* de $R_n$ l'ensemble des points $x$ dont les coordonnées satisfont aux inégalités $|x_j - x_j^0| < a$ ($j = 1, \ldots, n$).

Les coordonnées $x_j$ de tout point $x$ de la boule ouverte $|x - x^0| < r$ vérifient l'inégalité

$$|x_j-x_j^0| \leqslant \left(\sum_{i=1}^{n} (x_i-x_i^0)^2\right)^{1/2} < r.$$

Ce qui prouve que le point $x$ appartient au cube

$$|x_j - x_j^0| < r \quad (j = 1, \ldots, n).$$

Donc, *une boule de rayon r et de centre $x^0$ est contenue dans le cube de côté $2r$ et de centre $x^0$*. Ceci est illustré sur la figure 95 pour $n = 2$.

D'autre part, *le cube*

$$| x_j - x_j^0 | < a \quad (j = 1, \ldots, n)$$

*de côté* 2*a et de centre* $x^0$ *est contenu dans la boule de même centre et de rayon* $a\sqrt{n}$, car les coordonnées des points de ce cube vérifient l'inégalité

$$\Big(\sum_{j=1}^{n} (x_j - x_j^0)^2\Big)^{1/2} < \Big(\sum_{j=1}^{n} a^2\Big)^{1/2} = a\sqrt{n}.$$

Ceci est illustré sur la figure 96 pour $n = 2$.

Ce qui vient d'être dit pour les boules et cubes ouverts est valable pour les boules et cubes fermés.

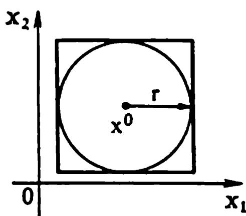



Fig. 95

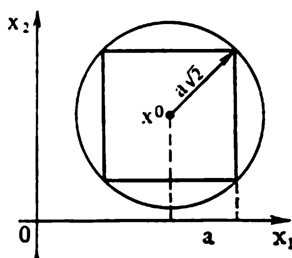



Fig. 96

Soit un ensemble $E$ de points $x$ de $R_n$. Par définition, $x^0$ est un *point intérieur* à l'ensemble $E$ s'il existe une boule ouverte de centre $x^0$ entièrement contenue dans $E$. On peut remplacer ici le terme *boule* par *cube*, car toute boule contient un cube de même centre et réciproquement.

Un ensemble est un *ouvert* si tous ses points sont intérieurs. Cette définition s'énonce encore comme suit: *un ensemble E est un ouvert si tout point est intérieur du seul fait d'appartenir à E.* Il s'ensuit qu'*un ensemble vide est un ouvert.*

*La boule ouverte*

$$| x - x^0 | < r \tag{1}$$

*est un ouvert.* En effet, soit $y$ un point de cette boule: $| y - x^0 | = \rho < r$. Pour tout point $x$ de la boule

$$| x - y | < \varepsilon \quad (\varepsilon < r - \rho) \tag{2}$$

on a

$$| x - x^0 | = | x - y + y - x^0 | \leqslant | x - y | + | y - x^0 | < \varepsilon + \rho < r.$$

Ceci montre que la boule (2) est contenue dans la boule (1). Donc, $y$ est un point intérieur et la boule (1), un ouvert.

On laisse au lecteur le soin de prouver qu'*un parallélotope ouvert, et notamment un cube ouvert, est un ouvert.*

On appelle *voisinage d'un point* $x^0 \in R_n$ un ouvert contenant $x^0$. Il est évident que *l'intersection de deux voisinages d'un point* $x^0$ *est encore un voisinage de* $x^0$.

D'après ce qui vient d'être dit on peut définir un point intérieur comme suit: *on dit qu'un point* $x^0$ *est un point intérieur d'un ensemble* $E$ *s'il existe un voisinage de* $x^0$ *contenu dans* $E$. En effet, si $x^0$ est un point intérieur au sens de la première définition, il existe une boule ouverte de centre $x^0$ contenue dans $E$, c'est-à-dire un voisinage de $x^0$. Réciproquement, si $x^0$ est un point intérieur au sens de la deuxième définition, il existe un voisinage de $x^0$ contenu dans $E$ et contenant en tant qu'ouvert une boule ouverte de centre $x^0$.

On étudiera dans la suite des fonctions $f(x) = f(x_1, \ldots, x_n)$ de $n$ variables $x_1, \ldots, x_n$ ou, ce qui revient au même, d'un point $x = (x_1, \ldots, x_n)$, définies sur des ouverts d'un espace à $n$ dimensions.

On dit qu'un ensemble $E$ est *connexe* s'il contient deux quelconques de ses points $x'$ et $x''$ avec une courbe continue les reliant, c'est-à-dire s'il existe une fonction vectorielle continue $x = x(t)$, $t \in$ $\in [0, 1]$ telle que $x(0) = x'$, $x(1) = x''$, $x(t) \in E$ (cf. § 1).

On appelle *segment* $[x', x'']$ d'origine $x'$ et d'extrémité $x''$ la portion de courbe $x(t) = tx' + (1 - t)x''$, $t \in [0, 1]$.

On appelle *domaine* un ouvert connexe.

## § 3. Limite d'une fonction

Au § 2 du chap. 3, nous avons étudié la notion de limite d'une fonction d'une variable. On se propose d'étendre cette notion aux fonctions de plusieurs variables.

On dit qu'*une fonction* $f(x) = f(x_1, \ldots, x_n)$ *a une limite* $A$ *au point* $x^0 = (x_1^0, \ldots, x_n^0)$ et on écrit

$$\lim_{x \to x^0} f(x) = \lim_{\substack{x_j \to x_j^0 \\ (j=1, \ldots, n)}} f(x_1, \ldots, x_n) = A \tag{1}$$

(ou encore $f(x) \to A$, $x \to x^0$) si elle est définie dans un voisinage de $x^0$, sauf éventuellement en $x^0$, et si existe la limite

$$\lim_{\substack{|x^k - x^0| \to 0 \\ x^k \neq x^0}} f(x^k) = A, \tag{2}$$

quelle que soit la suite de points $x^k \neq x^0$ de ce voisinage, convergeant vers $x^0$ (cf. § 1).

Voici une définition équivalente: on dit qu'*une fonction* $f$ *possède une limite* $A$ *en un point* $x^0$ si elle est définie dans un voisinage de $x^0$, sauf éventuellement en $x^0$, et si pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que

$$|f(x) - A| < \varepsilon \tag{3}$$

pour tous les $x$ vérifiant la double inégalité

$$0 < | x - x^0 | < \delta. \tag{4}$$

Cette définition équivaut à son tour à la suivante: pour tout $\varepsilon > 0$ on peut exhiber un voisinage $V(x^0)$ de $x^0$ tel que pour tous les $x \in V(x^0)$, $x \neq x^0$, on ait l'inégalité (3).

Formulons le critère de Cauchy qui se démontre comme en dimension un (cf. chap. 3, § 2, théorème 5).

*Pour qu'une fonction $f$ possède une limite (finie) en un point $x^0$, il est nécessaire et suffisant que pour tout $\varepsilon > 0$ il existe un voisinage $V(x^0)$* (notamment un cube ou une boule de centre $x^0$) *tel que pour tous les $x$, $x' \in V(x^0)$ distincts de $x^0$, l'on ait*

$$| f(x) - f(x') | < \varepsilon.$$

Si $A$ est la limite de $f(x)$ en $x^0$, il est alors évident qu'il est aussi limite de $f(x^0 + h)$ en $h = 0$:

$$\lim_{h \to 0} f(x^0 + h) = A,$$

et réciproquement.

Soient $f$ une fonction définie en tous les points de $V(x^0)$, sauf éventuellement en $x^0$, $\omega = (\omega_1, \ldots, \omega_n)$ un vecteur unitaire ($|\omega| = 1$), $t \geqslant 0$, un scalaire. Les points $x^0 + t\omega$ $(0 \leqslant t)$ forment une *demi-droite* issue de $x^0$ et de *vecteur directeur* $\omega$. Pour tout vecteur $\omega$ on peut considérer la fonction

$$f(x^0 + t\omega) = f(x_1^0 + t\omega_1, \ldots, x_n^0 + t\omega_n) \quad (0 < t < \delta_\omega)$$

de la variable scalaire $t$, où $\delta_\omega$ est un nombre dépendant de $\omega$. La limite (si elle existe) de cette fonction:

$$\lim_{\substack{t \to 0 \\ t > 0}} f(x^0 + t\omega) = \lim_{\substack{t \to 0 \\ t > 0}} f(x_1^0 + t\omega_1, \ldots, x_n^0 + t\omega_n)$$

s'appelle *limite de la fonction $f$ au point $x^0$ suivant le vecteur $\omega$.*

EXEMPLE 1. Soient les fonctions

$$f(x_1, x_2) = \frac{x_1^3 + x_2^3}{x_1^2 + x_2^2} \quad \text{et} \quad \varphi(x_1, x_2) = \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2}.$$

Elles sont définies sur le plan $(Ox_1, Ox_2)$ sauf en $x^0 = (0, 0)$. On a (compte tenu de ce que $| x_1 |^3 \leqslant (x_1^2 + x_2^2)^{3/2}$)

$$| f(x_1, x_2) | \leqslant \frac{2(x_1^2 + x_2^2)^{3/2}}{x_1^2 + x_2^2} = 2(x_1^2 + x_2^2)^{1/2} = 2 | x | = 2 | x - x^0 |,$$

d'où

$$\lim_{x_1, x_2 \to 0} f(x_1, x_2) = 0$$

(pour $\varepsilon > 0$ on admet que $\delta = \varepsilon/2$, et alors $| f(x_1, x_2) | < \varepsilon$ dès que $| x | < \delta$).

Si par ailleurs on admet que $k$ est une constante, alors $x_2 = kx_1$ et

$$\varphi(x_1, kx_1) = \frac{1-k^2}{1+k^2},$$

d'où il résulte que les limites de $\varphi$ en (0, 0) suivant des vecteurs différents ne sont pas égales. Donc, $\varphi$ ne possède pas de limite en (0, 0).

EXEMPLE 2. Considérons dans $R_2$ la fonction

$$f(x, y) = x^2y/(x^4 + y^2), \quad x^4 + y^2 \neq 0.$$

Cette fonction admet une limite nulle en (0, 0) suivant le vecteur directeur de toute droite $y = kx$ passant par l'origine:

$$f(x, kx) = kx^3/(x^4 + k^2x^2) = kx/(x^2 + k^2) \to 0 \quad \text{pour} \quad x \to 0.$$

Toutefois, cette fonction ne possède pas de limite au point (0, 0) puisque pour $y = x^2$

$$f(x, x^2) = x^4/(x^4 + x^4) = 1/2 \quad \text{et} \quad \lim_{\substack{x\to 0\\ y=x^2}} f(x, x^2) = 1/2.$$

On écrira $\lim\limits_{x\to x^0} f(x) = \infty$ si la fonction $f$ est définie dans un voisinage de $x^0$, sauf éventuellement en $x^0$, et si pour tout $N > 0$ on peut exhiber un $\delta > 0$ tel que $|f(x)| > N$ dès que $0 < |x - x^0| < \delta$.

On peut envisager la limite de $f$ pour $x \to \infty$:

$$\lim_{x\to\infty} f(x) = A. \tag{5}$$

Si le nombre $A$ est fini, il faut comprendre l'égalité (5) comme suit: pour tout $\varepsilon > 0$ on peut exhiber un $N > 0$ tel que pour tous les points $x$ vérifiant la relation $|x| > N$, la fonction $f$ est définie et l'on a $|f(x) - A| < \varepsilon$.

On a

$$\lim_{x\to x^0} (f(x) \pm \varphi(x)) = \lim_{x\to x^0} f(x) \pm \lim_{x\to x^0} \varphi(x), \tag{6}$$

$$\lim_{x\to x^0} (f(x)\cdot\varphi(x)) = \lim_{x\to x^0} f(x) \lim_{x\to x^0} \varphi(x), \tag{7}$$

$$\lim_{x\to x^0} \frac{f(x)}{\varphi(x)} = \frac{\lim\limits_{x\to x^0} f(x)}{\lim\limits_{x\to x^0} \varphi(x)} \quad (\lim_{x\to x^0} \varphi(x) \neq 0), \tag{8}$$

où éventuellement $x^0 = \infty$. Il est entendu que les limites (finies) des premiers membres existent s'il existe les limites de $f$ et de $\varphi$. Prouvons (7) à titre d'exemple.

Supposons que $x^k \to x^0$ $(x^k \neq x^0)$; alors

$$\lim_{x^k\to x^0} (f(x^k)\,\varphi(x^k)) = \lim_{x^k\to x^0} f(x^k)\cdot \lim_{x^k\to x^0} \varphi(x^k) = \lim_{x\to x^0} f(x) \lim_{x\to x^0} \varphi(x). \tag{9}$$

Donc, la limite du premier membre de (9) existe et est égale à celle du second membre. Comme la suite $\{x^k\}$ est arbitraire, cette limite est égale à celle de la fonction $f(x)\,\varphi(x)$ au point $x^0$.

THÉORÈME. *Si une fonction $f$ possède une limite non nulle en un point $x^0$:*

$$\lim_{x\to x^0} f(x) = A \neq 0.$$

*il existe un $\delta > 0$ tel que pour tous les $x$ vérifiant les inégalités*

$$0 < |x - x^0| < \delta, \tag{10}$$

*l'ont ait*

$$|f(x)| > |A|/2. \tag{11}$$

*En outre, la fonction $f$ garde le signe de $A$ pour les $x$ vérifiant* (10).

En effet, posons $\varepsilon = |A|/2$ et trouvons un $\delta > 0$ tel que pour tous les $x$ vérifiant les inégalités (10), l'on ait

$$|f(x) - A| < |A|/2. \tag{12}$$

Donc, pour ces $x$

$$|A|/2 > |A - f(x)| \geqslant |A| - |f(x)|,$$

c'est-à-dire qu'on obtient (11). De (12) il s'ensuit pour ces $x$:

$$A - \frac{|A|}{2} < f(x) < A + \frac{|A|}{2},$$

d'où $A/2 < f(x)$ pour $A > 0$ et $f(x) < A/2$ pour $A < 0$.

REMARQUE. Au § 12 de ce chapitre, on donnera une définition plus générale de la limite d'une fonction définie sur un ensemble arbitraire.

## § 4. Fonction continue

Par définition, *une fonction $f(x) = f(x_1, \ldots, x_n)$ est continue en un point $x^0 = (x_1^0, \ldots, x_n^0)$* si elle est définie dans un voisinage de $x^0$, y compris en $x^0$, et si sa limite en $x^0$ est égale à sa valeur en ce point:

$$\lim_{x\to x^0} f(x) = f(x^0). \tag{1}$$

La condition de continuité de $f$ en $x^0$ s'écrit sous la forme équivalente:

$$\lim_{h\to 0} f(x^0 + h) = f(x^0), \tag{1'}$$

autrement dit, la fonction $f(x)$ est continue en $x^0$ si la fonction $f(x^0 + h)$ l'est en $h = 0$.

On peut introduire l'accroissement de $f$ en $x^0$ correspondant à l'accroissement de $h = (h_1, \ldots, h_n)$:

$$\Delta_h f(x^0) = f(x^0 + h) - f(x^0)$$

et définir la continuité de $f$ en $x^0$ : *une fonction $f$ est continue en $x^0$ si*

$$\lim_{h\to 0} \Delta_h f(x^0) = \lim_{h_1, \ldots, h_n \to 0} [f(x_1^0 + h_1, \ldots, x_n^0 + h_n) - $$

$$- f(x_1^0, \ldots, x_n^0)] = 0. \quad (1'')$$

L'accroissement $\Delta_h f(x^0)$ s'appelle *accroissement total* de $f$ en $x^0$.

Des formules (6) à (8) du § 3 il s'ensuit immédiatement le

THÉORÈME 1. *La somme, la différence, le produit et le quotient de fonctions $f(x)$ et $\varphi(x)$ continues en $x^0$ est une fonction continue en ce point, pourvu que $\varphi(x^0) \neq 0$ dans le cas du quotient.*

Une constante $c$ peut être traitée comme une fonction $f(x) = c$ de $x = (x_1, \ldots, x_n)$. Cette fonction est continue quel que soit $x$, car

$$f(x + h) - f(x) = c - c = 0 \to 0 \quad (h \to 0).$$

La fonction $f_j(x) = x_j$, $j = 1, \ldots, n$, est également continue. En effet, soit $h = (h_1, \ldots, h_n)$; alors

$$|f_j(x + h) - f_j(x)| = |(x_j + h_j) - x_j| = |h_j| \leqslant |h| \to 0 \quad (h \to 0).$$

Si l'on effectue un nombre fini d'additions, de soustractions et de multiplications sur les fonctions $x_j$ et les constantes, on obtient des fonctions appelées *polynômes de $x$ ou de* $(x_1, \ldots, x_n)$. D'après les propriétés énumérées ci-dessus, *les polynômes sont des fonctions continues* sur $R_n$ (pour tous les $x \in R_n$). Le quotient $P/Q$ de deux polynômes est une *fonction* (ou *fraction*) *rationnelle* qui, de toute évidence, est continue sur $R_n$, sauf aux points $x$, où $Q(x) = 0$.

La fonction

$$P(x) = x_1^3 - x_2^3 + x_1^2 x_3 + 2x_1^2 x_2 - 3x_3^2 + 4$$

est un exemple de polynôme de $(x_1, x_2, x_3)$ du troisième degré.

On a le

THÉORÈME 2. *Soient $f(x) = f(x_1, \ldots, x_m)$ une fonction continue en un point $x^0 = (x_1^0, \ldots, x_m^0) \in R_m$, $\varphi_j(u) = \varphi_j(u_1, \ldots, u_n)$ des fonctions continues en un point $u^0 = (u_1^0, \ldots, u_n^0) \in R_n$. Si $\varphi_j(u^0) = x_j^0$ $(j = 1, \ldots, m)$, alors la fonction*

$$F(u) = f(\varphi_1(u), \varphi_2(u), \ldots, \varphi_m(u))$$

*est continue en $u^0$.*

DÉMONSTRATION. La fonction $f$ étant continue en $x^0$, pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que $f$ soit définie pour tous les $x$ pour lesquels $|x_j - x_j^0| < \delta$ $(j = 1, \ldots, m)$, et $|f(x) - f(x^0)| < \varepsilon$. Les fonctions $\varphi_j$ étant continues en $u^0 \in R_n$, on

peut indiquer un $\eta > 0$ tel que pour tous les points $u \in R_n$ de la boule $|u - u^0| < \eta$ l'on ait

$$|\varphi_j(u) - \varphi_j(u^0)| < \delta \quad (j = 1, \ldots, m).$$

On a alors

$$|F(u) - F(u^0)| =$$
$$= |f(\varphi_1(u), \ldots, \varphi_m(u)) - f(\varphi_1(u^0), \ldots, \varphi_m(u^0))| < \varepsilon,$$

c.q.f.d.

On appellera *fonction élémentaire* de $x_1, \ldots, x_n$ une fonction obtenue à partir de ces variables et de constantes $c$ par un nombre

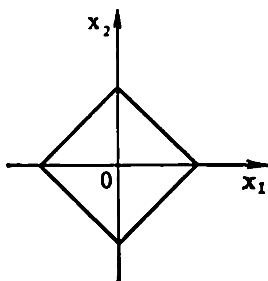



Fig. 97

fini d'additions, de soustractions, de multiplications, de divisions et d'opérations $\varphi$, où $\varphi$ sont des fonctions élémentaires d'une variable (cf. chap. 3, § 8). Comme exemples de fonctions élémentaires, citons

1) $\sin \ln \sqrt{1 + x^2 + y^2} = f_1$,

2) $\sin^2 x + \cos^3 (x + y) = f_2$, 3) $\ln \frac{x-y}{x+y} = f_3$.

Il est immédiat de vérifier en se servant des théorèmes 1 et 2 que les fonctions $f_1$ et $f_2$ sont continues sur le plan $(x, y)$, que $f_3$ est définie et continue aux points $(x, y)$, où la fraction $\frac{x-y}{x+y}$ est définie et positive.

Le théorème du § 3 et la définition de la continuité d'une fonction en un point entraînent immédiatement le

THÉORÈME 3. *Une fonction $f(x) = f(x_1, \ldots, x_n)$ continue et non nulle en $x^0$ conserve le signe de $f(x^0)$ dans un voisinage de $x^0$.*

COROLLAIRE. *Supposons qu'une fonction $f(x)$ est définie et continue sur $R_n$. Alors l'ensemble $G$ des points $x$, où elle vérifie l'inégalité $f(x) > c$ (ou $f(x) < c$), $\forall c$, est un ouvert.*

En effet, la fonction $F(x) = f(x) - c$ est continue sur $R_n$ et l'ensemble de tous les points $x$ tels que $F(x) > 0$ est confondu avec $G$. Soit $x^0 \in G$; il existe alors une boule

$$|x - x^0| < \delta$$

sur laquelle $F(x) > 0$, c'est-à-dire que cette boule est contenue dans $G$ et le point $x^0 \in G$ est intérieur à $G$.

Le cas $f(x) < c$ se démontre de façon analogue.

EXEMPLE. Les fonctions

$$f_1(x) = \sum_1^n \frac{x_k^2}{a_k} \quad (a_k > 0);$$

$$f_2(x) = \sum_1^n |x_k|,$$

sont définies et continues sur $R_n$.

Dans ce cas, les ensembles des valeurs de $x$ telles que $f_i(x) < c$ ($i = 1, 2$) sont des ouverts. Le premier est un ellipsoïde ouvert d'un espace à $n$ dimensions, le second, pour $n = 2$, un carré ouvert représenté sur la figure 97.

## § 5. Dérivées partielles et dérivée suivant un vecteur

Dans ce paragraphe, on se propose d'étudier des fonctions $f$ définies sur un ouvert $G \subset R_n$.

On appellera *accroissement d'une fonction $f$ de pas $h$ en un point $x = (x_1, \ldots, x_n) \in G$ par rapport à la variable $x_j$* la quantité

$$\Delta_{x_j h} f(x) =$$
$$= f(x_1, \ldots, x_{j-1}, x_j + h, x_{j+1}, \ldots, x_n) - f(x_1, \ldots, x_n),$$

où $h$ est un réel assez petit pour que cette quantité ait un sens.

On appelle *dérivée partielle de $f$ par rapport à $x_j$ en un point $x$* la limite (si elle existe)

$$f'_{x_j} = \frac{\partial f}{\partial x_j} = \frac{\partial f(x)}{\partial x_j} = \lim_{h \to 0} \frac{\Delta_{x_j h} f(x)}{h} \quad (j = 1, \ldots, n).$$

La dérivée partielle $\frac{\partial f(x)}{\partial x_j}$ est la dérivée ordinaire de la fonction $f(x_1, \ldots, x_n)$ traitée comme une fonction de la variable $x_j$ à $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ fixes.

Une fonction $z = f(x, y)$ de deux variables est représentée graphiquement dans l'espace par une surface. Il est évident que $f'_x(x_0, y_0)$ (si elle existe) *est égale à la tangente de l'angle entre l'axe $Ox$ et la tangente au point d'abscisse $x_0$ à la section de cette surface par le plan $y = y_0$.*

Si une fonction $u = f(x_1, \ldots, x_n)$ possède des dérivées partielles $\frac{\partial f}{\partial x_k}$ ($k = 1, \ldots, n$) en tous les points $x \in G$, alors ces dérivées peuvent être traitées comme des fonctions définies sur $G$.

On peut donc poser le problème de l'existence des dérivées partielles de ces fonctions par rapport à une variable en $x$.

Si la fonction $\frac{\partial f}{\partial x_k}$ possède une dérivée partielle par rapport

à $x_k$, on appelle cette dernière, *dérivée partielle seconde de la fonction* $f(x)$ *par rapport à* $x_k$ et on la note $\frac{\partial^2 f}{\partial x_k^2}$. Donc, par définition

$$\frac{\partial^2 f}{\partial x_k^2} = \frac{\partial}{\partial x_k}\left\{\frac{\partial f}{\partial x_k}\right\}.$$

La dérivée partielle de la fonction $\frac{\partial f}{\partial x_k}$ par rapport à une autre variable $x_i$ $(i \neq k)$ s'appelle *dérivée partielle mixte d'ordre deux* et se note

$$\frac{\partial^2 f}{\partial x_i\, \partial x_k} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_k}\right).$$

D'une façon générale, on appelle *dérivée partielle d'ordre n* la dérivée partielle d'une dérivée d'ordre $n-1$ par rapport à une variable. Exemple:

$$\frac{\partial}{\partial x}\left(\frac{\partial^3 f}{\partial x^2\, \partial y}\right) = \frac{\partial^4 f}{\partial x^3\, \partial y}.$$

Les dérivées partielles $\frac{\partial f}{\partial x_k}$ seront dites *dérivées partielles premières* ou *d'ordre un* et la fonction $f$, *dérivée partielle d'ordre zéro*.

Pour désigner les dérivées partielles, on se servira encore des notations:

$$D_{x_1} f = \frac{\partial f}{\partial x_1}, \quad D_{x_1} D_{x_2} f = \frac{\partial^2 f}{\partial x_1\, \partial x_2}, \quad \ldots, \quad \frac{\partial^2 f}{\partial x^2} = f''_{x^2}, \quad \frac{\partial^2 f}{\partial x\, \partial y} = f''_{xy}.$$

Si $r = (r_1, \ldots, r_n)$ est un vecteur dont les coordonnées sont entières et positives, on écrira

$$D^r f = D_{x_1}^{r_1} \ldots D_{x_n}^{r_n} f = \frac{\partial^{r_1 + \ldots + r_n} f}{\partial x_1^{r_1} \ldots \partial x_n^{r_n}}.$$

EXEMPLE. Calculer la dérivée $\frac{\partial^3 f}{\partial x\, \partial y^2}$ de la fonction $f(x, y) = x^2 + \sin xy$. On a

$$\frac{\partial f}{\partial y} = x \cos xy, \quad \frac{\partial^2 f}{\partial y^2} = -x^2 \sin xy,$$

$$\frac{\partial^3 f}{\partial x\, \partial y^2} = -2x \sin xy - x^2 y \cos xy.$$

On peut poser tout naturellement le problème de l'égalité des dérivées partielles prises par rapport aux mêmes variables, le même nombre de fois, mais dans un ordre différent.

Par exemple, est-ce que les dérivées

$$\frac{\partial^4 f}{\partial x\, \partial y^2\, \partial x} \quad \text{et} \quad \frac{\partial^4 f}{\partial x^2\, \partial y^2}$$

sont égales?

D'une façon générale, la réponse est négative. On a cependant le théorème suivant.

THÉORÈME (DES DÉRIVÉES MIXTES). *Si une fonction* $u = f(x, y)$ *est définie avec ses dérivées partielles* $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial^2 f}{\partial x\, \partial y}$, $\frac{\partial^2 f}{\partial y\, \partial x}$ *dans un voisinage d'un point* $P_0 = (x_0, y_0)$ *et de plus* $\frac{\partial^2 f}{\partial x\, \partial y}$ *et* $\frac{\partial^2 f}{\partial y\, \partial x}$ *sont continues en* $P_0$, *alors*

$$\frac{\partial^2 f(P_0)}{\partial x\, \partial y} = \frac{\partial^2 f(P_0)}{\partial y\, \partial x},$$

*autrement dit, la dérivation ne dépend pas de l'ordre dans lequel elle a été effectuée.*

Nous passerons sur la démonstration de ce théorème et nous nous bornerons aux remarques suivantes.

REMARQUE 1. On peut généraliser aisément par récurrence ce théorème à des dérivées partielles mixtes continues quelconques, ne différant entre elles que par l'ordre de la dérivation. Par exemple,

$$\frac{\partial^3 f}{\partial x\, \partial y^2} = \frac{\partial}{\partial x}\left\{\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial y}\right]\right\} = \frac{\partial}{\partial y}\left\{\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right]\right\} \equiv$$

$$\equiv \frac{\partial^3 f}{\partial y\, \partial x\, \partial y} = \frac{\partial}{\partial y}\left\{\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\right]\right\} \equiv \frac{\partial^3 f}{\partial y^2\, \partial x}.$$

REMARQUE 2. Si la continuité n'est pas assurée, les dérivées mixtes peuvent être différentes en $P_0$. Soit la fonction

$$f(x, y) = \frac{x^3 y}{x^2 + y^2} \quad (x^2 + y^2 \neq 0,\ f(0, 0) = 0).$$

On a aussitôt

$$f'_x(x, y) = y\,\frac{x^4 + 3x^2y^2}{(x^2+y^2)^2} \quad \text{et} \quad f'_x(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x} = 0;$$

$$f'_y(x, y) = x^3\,\frac{x^2 - y^2}{(x^2+y^2)^2} \quad \text{et} \quad f'_y(0, 0) = 0.$$

Par définition,

$$f''_{xy}(0, 0) = \lim_{x \to 0} \frac{f'_y(x, 0) - f'_y(0, 0)}{x} = \lim_{x \to 0} \frac{x - 0}{x} = 1,$$

$$f''_{yx}(0, 0) = \lim_{y \to 0} \frac{f'_x(0, y) - f'_x(0, 0)}{y} = \lim_{y \to 0} \frac{0 - 0}{y} = 0,$$

c'est-à-dire que

$$f''_{xy}(0, 0) \neq f''_{yx}(0, 0).$$

Signalons que les dérivées partielles $f''_{xy}$ et $f''_{yx}$ sont discontinues au point (0, 0). En effet,

$$f''_{yx}(x, y) = \frac{x^6 + 6x^4y^2 - 3x^2y^4}{(x^2+y^2)^3},$$
$$x^2 + y^2 \neq 0,$$

d'où

$$\lim_{\substack{x \to 0 \\ x=y}} f''_{yx}(x, y) = \frac{1}{2} \neq f''_{yx}(0, 0) = 0.$$

Introduisons la notion de *dérivée suivant un vecteur*. Cette notion n'a pas de sens pour une fonction d'une variable.

Soit $\omega = (\omega_1, \ldots, \omega_n)$ un vecteur unitaire arbitraire. On appelle *dérivée d'une fonction $f$ suivant le vecteur $\omega$, en un point $x$*, la limite (si elle existe)

$$\frac{\partial f}{\partial \omega} = \lim_{\substack{t \to 0 \\ t > 0}} \frac{f(x + t\omega) - f(x)}{t}.$$

A noter que dans le calcul de cette limite on admet que $t$ tend vers 0 par valeurs strictement *positives*, donc on peut encore dire que $\frac{\partial f(x)}{\partial \omega}$ *est la dérivée à droite par rapport à $t$ de la fonction $f(x + t\omega)$ au point $t = 0$.*

On peut, comme dans le cas de fonctions d'une variable, parler de dérivées partielles à droite et à gauche par rapport à $x_j$. A noter simplement que *la dérivée suivant le vecteur $i_j$ de l'axe des $x_j$ est confondue avec la dérivée partielle à droite par rapport à $x_j$, mais que la dérivée suivant le vecteur $-i_j$ est du signe contraire de celui de la dérivée à gauche par rapport à $x_j$.*

## § 6. Fonctions différentiables

On se placera dans un espace à trois dimensions par souci de simplicité. Les raisonnements sont identiques dans un espace à $n$ dimensions. Le cas $n = 1$ a été spécialement étudié au § 7 du chapitre 4.

Soit donnée sur un ouvert $G \subset R_3$ une fonction $u = f(x, y, z)$ possédant en un point $(x, y, z)$ de $G$ des dérivées partielles continues du premier ordre. Il s'ensuit automatiquement que ces dérivées existent dans un voisinage de $(x, y, z)$, bien qu'éventuellement elles soient susceptibles de ne pas être continues aux points distincts de $(x, y, z)$.

Considérons l'accroissement de $f$ en $(x, y, z)$ correspondant à l'accroissement $(\Delta x, \Delta y, \Delta z)$, où $|\Delta x|$, $|\Delta y|$, $|\Delta z|$ sont strictement inférieurs à un nombre $\delta$ assez petit pour que le point

$(x + \Delta x, y + \Delta y, z + \Delta z)$ soit contenu dans le voisinage indiqué. On a les égalités suivantes:

$$\begin{aligned}
\Delta u &= f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z) = && (1)\\
&= f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y + \Delta y, z + \Delta z) + && (2)\\
&+ f(x, y + \Delta y, z + \Delta z) - f(x, y, z + \Delta z) + && (3)\\
&\qquad + f(x, y, z + \Delta z) - f(x, y, z) = && (4)\\
&= f'_x(x + \theta_1 \Delta x, y + \Delta y, z + \Delta z)\, \Delta x + \\
&+ f'_y(x, y + \theta_2 \Delta y, z + \Delta z)\, \Delta y + f'_z(x, y, z + \theta_3 \Delta z)\, \Delta z = && (5)\\
&= (f'_x(x, y, z) + \varepsilon_1)\, \Delta x + (f'_y(x, y, z) + \varepsilon_2)\, \Delta y + \\
&\qquad + (f'_z(x, y, z) + \varepsilon_3)\, \Delta z = && (6)\\
&= f'_x(x, y, z)\, \Delta x + f'_y(x, y, z)\, \Delta y + f'_z(x, y, z)\, \Delta z + o(\rho) && (7)\\
&\qquad (\rho \to 0),
\end{aligned}$$

$$0 < \theta_1, \theta_2, \theta_3 < 1, \ \rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}, \ \varepsilon_1, \varepsilon_2, \varepsilon_3 \to 0 \quad (\rho \to 0). \qquad (8)$$

Le passage de (2) au premier terme de (5) se justifie de la manière suivante: à la fonction $f(\xi, y + \Delta y, z + \Delta z)$ de $\xi$, dérivable (à $y + \Delta y$, $z + \Delta z$ fixes) par rapport à $\xi$ sur $[x, x + \Delta x]$ par hypothèse, on a appliqué le théorème des accroissements finis. On a fait de même pour le second et le troisième terme de (5). Le passage de (5) à (6) est purement formel: nous avons posé par exemple:

$$f'_x(x + \theta_1 \Delta x, y + \Delta y, z + \Delta z) = f'_x(x, y, z) + \varepsilon_1.$$

Mais il n'est pas évident que $\varepsilon_1 \to 0$ avec $\rho$. Ce fait découle de la continuité (par hypothèse) de $f'_x$ en $(x, y, z)$. Enfin, le passage de (6) à (7) revient à affirmer que

$$\varepsilon_1 \Delta x + \varepsilon_2 \Delta y + \varepsilon_3 \Delta z = o(\rho) \quad (\rho \to 0).$$

En effet, puisque $|\Delta x|, |\Delta y|, |\Delta z| \leqslant \rho$, on a pour $\rho \to 0$

$$|\varepsilon_1 \Delta x + \varepsilon_2 \Delta y + \varepsilon_3 \Delta z|/\rho \leqslant |\varepsilon_1| + |\varepsilon_2| + |\varepsilon_3| \to 0.$$

Nous venons de prouver l'important théorème suivant.

THÉORÈME 1. *Si une fonction $u = f$ possède des dérivées partielles continues (du premier ordre) en un point $(x, y, z)$, alors son accroissement en ce point peut s'écrire*

$$\Delta u = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \frac{\partial f}{\partial z} \Delta z + o(\rho) \quad (\rho \to 0), \qquad (9)$$

$$\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2},$$

*où les dérivées partielles sont prises au point $(x, y, z)$.*

Comme les valeurs des dérivées partielles du second membre de (9) ne dépendent pas de $\Delta x$, $\Delta y$, $\Delta z$, le théorème 1 nous dit que l'accroissement de $f$ en $(x, y, z)$ peut s'écrire

$$\Delta u = A\, \Delta x + B\, \Delta y + C\, \Delta z + o(\rho) \quad (\rho \to 0), \qquad (10)$$

où les nombres $A$, $B$ et $C$ ne dépendent pas de $\Delta x$, $\Delta y$ et de $\Delta z$.

Introduisons la définition suivante: on dira qu'*une fonction f est différentiable en un point* $(x, y, z)$ si l'accroissement de $f$ en $(x, y, z)$ pour $(\Delta x, \Delta y, \Delta z)$ assez petits peut être mis sous la forme (10), où $A$, $B$ et $C$ sont des constantes ne dépendant pas de $\Delta x$, $\Delta y$ et de $\Delta z$. Donc, *dire qu'une fonction f est différentiable en* $(x, y, z)$ *revient à dire que son accroissement* $\Delta f$ *en* $(x, y, z)$ *peut être mis sous la forme d'une somme de deux termes: le premier est une fonction linéaire* $A\ \Delta x + B\ \Delta x + C\ \Delta z$ *de* $(\Delta x, \Delta y, \Delta z)$ *appelée partie principale de l'accroissement* $\Delta f$, *le second dépend d'une manière complexe des accroissements* $\Delta x$, $\Delta y$, $\Delta z$, *mais si ces derniers tendent vers* 0, *ce terme tendra vers* 0 *plus vite que* $\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$.

Il est immédiat de voir que si une fonction $f$ est différentiable en un point $(x, y, z)$, c'est-à-dire représentable par (10), alors elle possède en ce point des dérivées partielles du premier ordre égales à:

$$\frac{\partial f}{\partial x} = A, \quad \frac{\partial f}{\partial y} = B, \quad \frac{\partial f}{\partial z} = C. \tag{11}$$

La première égalité de (11), par exemple, se démontre de la manière suivante. Supposons que l'accroissement de $f$ en $(x, y, z)$ se représente par la formule (10). Si l'on pose $\Delta x = h$, $\Delta y = \Delta z = 0$, on obtient l'égalité $\Delta_{xh} u = Ah + o(h)$ $(h \to 0)$. Une division par $h$ et un passage à la limite nous donnent:

$$\lim_{h \to 0} \frac{\Delta_{xh} u}{h} = \frac{\partial f}{\partial x} = A.$$

De ce qui précède résulte le

THÉORÈME 2. *Pour qu'une fonction f soit différentiable en un point il est nécessaire qu'elle y possède des dérivées partielles et suffisant qu'elle y possède des dérivées partielles continues.*

On rappelle que dans le cas d'une fonction d'une variable, l'existence de la dérivée en $x$ est une condition nécessaire et suffisante de différentiabilité en $x$.

De (10) il s'ensuit que *si une fonction est différentiable en un point elle y est nécessairement continue.*

EXEMPLE 1. La fonction $f(x, y, z)$ nulle sur les plans de coordonnées $x = 0$, $y = 0$ et $z = 0$ et égale à un dans les autres points de $R_3$ possède de toute évidence des dérivées partielles nulles en (0, 0, 0), mais elle est discontinue en ce point, donc n'y est pas différentiable. Par conséquent, la seule existence des dérivées partielles en un point ne garantit pas la différentiabilité et encore moins la continuité en ce point.

Signalons la différence existant entre la dimension un et la dimension $n$. Pour $n = 1$, la différentiabilité de $f$ est traduite par l'égalité $\Delta f = A\ \Delta x + o(\Delta x)$, donc si $A \neq 0$, le reste tend vers 0 avec $\Delta x$ plus vite que la partie principale. Pour $n > 1$, il en va autrement. Ainsi, pour $n = 3$, quels que soient

$A$, $B$ et $C$ non tous nuls, on peut toujours faire tendre $\Delta x$, $\Delta y$ et $\Delta z$ vers 0 de telle sorte que l'on ait en permanence $A\ \Delta x + B\ \Delta y + C\ \Delta z = 0$, et alors dans (10), le reste $o(\rho)$ sera plus grand que le terme principal. Au demeurant, si l'on fait tendre $\Delta x$, $\Delta y$ et $\Delta z$ vers 0 de telle sorte que $\Delta x : \Delta y : \Delta z = A : B : C$, alors la partie principale de l'accroissement sera du même ordre que $\rho$, et le reste tendra vers 0 plus vite que la partie principale.

EXEMPLE 2. La fonction $u = |x|(y + 1)$ est continue au point (0, 0). Cependant, on voit immédiatement que $\frac{\partial u}{\partial x}$ n'existe pas en (0, 0). Donc, $u$ n'est pas différentiable en (0, 0).

Si une fonction $f$ est différentiable en un point $(x, y, z)$, la partie principale de son accroissement en ce point s'appelle *différentielle de $f$ en $(x, y, z)$ correspondant aux accroissements* $(\Delta x, \Delta y, \Delta z)$ *des variables indépendantes* et se note

$$df = \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \frac{\partial f}{\partial z} \Delta z.$$

Les autres notations seront examinées au § 9.

## § 7. Plan tangent. Interprétation géométrique de la différentielle

Soit donnée une surface $S$ d'équation

$$z = f(x, y), \tag{1}$$

où $f(x, y)$ possède des dérivées partielles continues sur un domaine du plan $(Ox, Oy)$. On peut admettre que $f$ est différentiable en tout point de ce domaine.

On appelle *plan tangent* à la surface $S$ en $M_0 = (x_0, y_0, z_0)$, le plan d'équation

$$Z - z_0 = \left(\frac{\partial f}{\partial x}\right)_0 (X - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (Y - y_0), \tag{2}$$

où $X$, $Y$ et $Z$ sont les coordonnées courantes, $\left(\frac{\partial f}{\partial x}\right)_0$ et $\left(\frac{\partial f}{\partial y}\right)_0$ les valeurs des dérivées partielles de $f$ au point $P_0 = (x_0, y_0)$.

Désignons le plan (2) par $\Pi$. Ce plan passe par $M_0$ et jouit d'une propriété qui le distingue des autres plans passant par $M_0$.

Soit $P = (x, y)$ un point du plan $(Ox, Oy)$ proche de $P_0 = (x_0, y_0)$ (fig. 98). La parallèle menée par $P$ à l'axe $Oz$ coupe $\Pi$ en $T$ et la surface $S$ en $M$. La cote de $M$ est

$$z = f(x, y),$$

celle de $T$ a pour valeur

$$Z = f(x_0, y_0) + \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0).$$

La distance des points $M$ et $T$ est égale à

$$|MT| = |f(x, y) - f(x_0, y_0) - \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) - \\ - \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0)|. \quad (3)$$

Celle des points $P$ et $P_0$ à

$$\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2}.$$

La fonction $f$ est différentiable en $(x_0, y_0)$, car elle y admet par hypothèse des dérivées partielles continues. Donc, le second membre



Fig. 98

de (3) tend vers 0 plus vite que $\rho$, c'est-à-dire que

$$|MT| = o(\rho) \quad (\rho \to 0).$$

Nous venons de prouver que *le plan* $\Pi$ *tangent à la surface* $S$ *en* $M_0 = (x_0, y_0, z_0)$ *passe par ce dernier et jouit de la propriété suivante : la distance suivant l'axe* $Oz$ *d'un point quelconque* $(x, y, f(x, y))$ *de la surface* $S$ *au plan* $\Pi$ *est égale à* $o(\rho)$ $(\rho \to 0)$, *où* $\rho$ *est la distance des points* $(x, y)$ *et* $(x_0, y_0)$ *du plan* $(Ox, Oy)$.

Cette propriété est caractéristique du plan tangent, car si un plan $\Pi'$ de la forme

$$Z - z_0 = a(x - x_0) + b(y - y_0) \quad (z_0 = f(x_0, y_0))$$

possède cette propriété, c est-à-dire s'il est justiciable de l'égalité

$$f(x, y) - f(x_0, y_0) - a(x - x_0) - b(y - y_0) = o(\rho) \quad (\rho \to 0)$$

ou ce qui revient au même, de l'égalité

$$f(x, y) - f(x_0, y_0) = a(x - x_0) + b(y - y_0) + o(\rho) \quad (\rho \to 0),$$

alors nous savons que $f$ est différentiable en $(x_0, y_0)$ et

$$a = \left(\frac{\partial f}{\partial x}\right)_0, \quad b = \left(\frac{\partial f}{\partial y}\right)_0,$$

autrement dit le plan $\Pi'$ est tangent à $S$ $(\Pi' = \Pi)$.

Donc, *pour qu'une surface $S$ admette un plan tangent en un point $(x_0, y_0, f(x_0, y_0))$ il est nécessaire et suffisant que la fonction $f$ soit différentiable au point $P_0 = (x_0, y_0)$.*

Le second membre de l'équation (2) est la différentielle de $f$ au point $(x_0, y_0)$:

$$df = \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0).$$

Le premier membre de (2) est l'accroissement correspondant de la cote du plan tangent $\Pi$.

Par conséquent, *du point de vue géométrique, la différentielle de la fonction $f$ en $(x_0, y_0)$ est l'accroissement de la cote du point $(x, y)$ du plan tangent à la surface $z = f(x, y)$ en $(x_0, y_0)$.*

REMARQUE. Si la fonction $z = f(x, y)$ n'est pas différentiable en $(x_0, y_0)$, bien qu'elle y admette des dérivées partielles, on ne peut pas dire que le plan (2) est tangent à la surface $z = f(x, y)$, car en ce point la différence $f(x, y) - Z$ ne tend pas vers 0 plus vite que $\rho$ pour $\rho \to 0$. Si par exemple la fonction $z = f(x, y)$ est nulle sur les axes $Ox$ et $Oy$ et égale à 1 dans les autres points du plan $(Ox, Oy)$, alors $f'_x(0, 0) = f'_y(0, 0) = 0$ et l'équation (2) devient $Z = 0$. La différence $f(x, y) - Z = f(x, y) - 0 = 1$ pour tous les points $(x, y)$ non situés sur les axes $Ox$ et $Oy$. Donc, cette différence ne tend même pas vers 0 pour $\rho \to 0$.

## § 8. Dérivée d'une fonction composée. Dérivée suivant un vecteur. Gradient. Fonctions homogènes

On se bornera à étudier des fonctions de trois variables définies sur un ouvert $G \subset R_3$. La généralisation à $R_n$ est immédiate.

THÉORÈME 1. *Soient*

$$u = f(x, y, z) \tag{1}$$

*une fonction différentiable en $(x, y, z) \in G$,*

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t), \tag{2}$$

*des fonctions d'un paramètre scalaire $t$, dérivables en tout point $t$. Alors la dérivée par rapport à $t$ (c'est-à-dire la dérivée le long de la courbe (2)) de la fonction composée $u = F(t) = f(\varphi(t), \psi(t), \chi(t))$*

*a pour expression*

$$F'(t) = f'_x(\varphi(t), \psi(t), \chi(t))\,\varphi'(t) + \\ + f'_y(\varphi(t), \psi(t), \chi(t))\,\psi'(t) + f'_z(\varphi(t), \psi(t), \chi(t))\,\chi'(t),$$

*ou de façon plus concise*

$$\frac{du}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}. \tag{3}$$

En effet, la fonction $f$ étant différentiable en $(x, y, z)$, quelque petit que soit l'accroissement $(\Delta x, \Delta y, \Delta z)$, on a

$$\Delta u = f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z) = \\ = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z + o(\rho) \tag{4}$$

$$\left(\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \to 0\right).$$

A un accroissement $\Delta t$ de $t$ correspondent les accroissements $\Delta x$, $\Delta y$ et $\Delta z$ des fonctions (2). En portant ces accroissements dans (4), on obtient l'accroissement $F(t + \Delta t) - F(t) = \Delta u$ de la fonction $F$ en $t$. Une division de (4) par $\Delta t$ et le passage à la limite nous donnent

$$F'(t) = \lim_{\Delta t \to 0} \frac{\Delta u}{\Delta t} = \\ = \lim_{\Delta t \to 0}\left(\frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t} + \frac{\partial f}{\partial z}\frac{\Delta z}{\Delta t} + \frac{o(\rho)}{\Delta t}\right) = \\ = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt},$$

c'est-à-dire (3), car les fonctions (2) sont dérivables et

$$\frac{o(\rho)}{\Delta t} = \varepsilon(\rho)\sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2 + \left(\frac{\Delta z}{\Delta t}\right)^2} \to \\ \to 0 \cdot \sqrt{x_t'^2 + y_t'^2 + z_t'^2} = 0 \quad (\Delta t \to 0)$$

($\Delta t \to 0$ implique $\rho \to 0$).

REMARQUE. Si les fonctions $x$, $y$, $z$ dépendent de plusieurs variables, par exemple de deux:

$$x = \varphi(t, \tau), \quad y = \psi(t, \tau), \quad z = \chi(t, \tau),$$

alors en fixant d'abord $\tau$ et ensuite $t$, on trouve en vertu de (3)

$$\frac{\partial u}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t},$$

$$\frac{\partial u}{\partial \tau} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \tau} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \tau} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial \tau}.$$

THÉORÈME 2. *Si une fonction f est différentiable en un point (x, y, z), alors la dérivée suivant tout vecteur unitaire* $\boldsymbol{n} = (\cos\alpha, \cos\beta, \cos\gamma)$, *exprimée par*

$$\frac{\partial f}{\partial n}=\frac{\partial f}{\partial x}\cos\alpha+\frac{\partial f}{\partial y}\cos\beta+\frac{\partial f}{\partial z}\cos\gamma \tag{5}$$

($\alpha$, $\beta$ *et* $\gamma$ *étant les angles de* $\boldsymbol{n}$ *avec les axes* $Ox$, $Oy$ *et* $Oz$), *a un sens pour elle.*

DÉMONSTRATION. Par définition de la dérivée suivant un vecteur (cf. § 5) et en vertu du théorème précédent, on a

$$\frac{\partial f}{\partial n}=\lim_{\substack{t\to 0\\ t>0}}\frac{f(x+t\cos\alpha,\ y+t\cos\beta,\ z+t\cos\gamma)-f(x,y,z)}{t}=$$

$$=\left[\frac{d}{dt}f(x+t\cos\alpha,\ y+t\cos\beta,\ z+t\cos\gamma)\right]_{t=0}=$$

$$=\frac{\partial f}{\partial x}\cos\alpha+\frac{\partial f}{\partial y}\cos\beta+\frac{\partial f}{\partial z}\cos\gamma,$$

où les dérivées partielles sont prises au point $(x, y, z)$.

Si $x=\varphi(s)$, $y=\psi(s)$, $z=\chi(s)$ sont les équations d'une courbe lisse $\Gamma$, où le paramètre $s$ représente la longueur d'arc, alors

$$\frac{dx}{ds}=\varphi'(s),\qquad \frac{dy}{ds}=\psi'(s),\qquad \frac{dz}{ds}=\chi'(s)$$

sont les cosinus directeurs du vecteur tangent à $\Gamma$. Donc,

$$\frac{\partial f}{\partial s}=\frac{\partial f}{\partial x}\varphi'(s)+\frac{\partial f}{\partial y}\psi'(s)+\frac{\partial f}{\partial z}\chi'(s)=\frac{d}{ds}f(\varphi(s),\ \psi(s),\ \chi(s)),$$

où $f$ est une fonction différentiable, est la dérivée suivant le vecteur tangent indiqué. On dit encore que $\frac{\partial f}{\partial s}$ est la dérivée de $f$ le long de $\Gamma$.

Considérons le vecteur

$$\operatorname{grad} f=\left(\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y},\ \frac{\partial f}{\partial z}\right) \tag{6}$$

appelé *gradient* de la fonction $f$ au point $(x, y, z)$.

La formule (5) dit que *la dérivée de f en (x, y, z) suivant le vecteur unitaire* $\boldsymbol{n}$ *est égale* à *la projection du gradient en ce point sur la direction de* $\boldsymbol{n}$:

$$\frac{\partial f}{\partial n}=(\operatorname{grad} f,\ \boldsymbol{n})=\operatorname{grad}_n f. \tag{7}$$

On a l'inégalité évidente

$$\frac{\partial f}{\partial n}\leqslant|\operatorname{grad} f| \tag{8}$$

pour tout vecteur $\boldsymbol{n}$. Si grad $f = 0$ (ceci n'a lieu qu'en des points exceptionnels), alors $\frac{\partial f}{\partial n} = 0$ pour tout $\boldsymbol{n}$. Si grad $f \neq 0$ (l'une des dérivées partielles de $f$ n'est pas nulle), alors (8) est une inégalité stricte pour tous les vecteurs unitaires $\boldsymbol{n}$, à l'exception du vecteur unitaire $\boldsymbol{n}_0 = (\cos\alpha_0, \cos\beta_0, \cos\gamma_0)$ dirigé dans le sens de grad $f$ $\left(\frac{\partial f}{\partial \boldsymbol{n}_0} = |\text{grad } f| > 0\right)$. Donc

$$\cos\alpha_0 = \frac{\frac{\partial f}{\partial x}}{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}},$$
$$\cos\beta_0 = \frac{\frac{\partial f}{\partial y}}{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}}, \tag{9}$$
$$\cos\gamma_0 = \frac{\frac{\partial f}{\partial z}}{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2}}.$$

De ce qui précède il suit que *le gradient d'une fonction $f$ en un point $(x, y, z)$ peut se définir comme un vecteur doué des deux propriétés suivantes*:

1) *son module est égal à la valeur maximale de la dérivée* $\frac{\partial f}{\partial \boldsymbol{n}}$ *en* $(x, y, z)$ (ce maximum existe et est positif pour toute fonction différentiable en $(x, y, z)$);

2) *s'il n'est pas nul, il est de même sens que le vecteur* $\boldsymbol{n}$ *suivant lequel la dérivée* $\frac{\partial f}{\partial n}$ *est maximale.*

Considérons les fonctions dites *homogènes*.

Soit donné un vecteur $\lambda = (\lambda_1, \ldots, \lambda_n)$, où $\lambda_i$ sont des nombres arbitraires. Une fonction $f(x_1, \ldots, x_n)$ définie sur $R_n$ est dite *$\lambda$-homogène de degré $m$* si pour tout $t > 0$ et pour tout $\boldsymbol{x} = (x_1, \ldots, x_n) \in R_n$, on a

$$f(t^{\lambda_1}x_1, t^{\lambda_2}x_2, \ldots, t^{\lambda_n}x_n) = t^{m\frac{|\lambda|}{n}} f(x_1, \ldots, x_n), \tag{10}$$

où $|\lambda| = \sum_{i=1}^{n} |\lambda_i|$.

Si $\lambda_1 = \lambda_2 = \ldots = \lambda_n$, la fonction $f$ s'appelle simplement *fonction homogène de degré $m$.*

THÉORÈME 3. *Pour qu'une fonction $f$ possédant des dérivées partielles du premier ordre continues soit $\lambda$-homogène de degré $m$, il est nécessaire*

*et suffisant que l'on ait*

$$\sum_{i=1}^{n} \lambda_i x_i f'_{x_i}(x_1, \ldots, x_n) = m \frac{|\lambda|}{n} f(x_1, \ldots, x_n). \tag{11}$$

Si la fonction $f$ est homogène de degré $m$, on obtient le théorème d'Euler.

DÉMONSTRATION 1). Supposons que la fonction $f$ est $\lambda$-homogène de degré $m$. En dérivant (10) par rapport à $t$, on trouve

$$\sum_{i=1}^{n} \frac{\partial f(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n)}{\partial x_i} \lambda_i x_i t^{\lambda_i - 1} = m \frac{|\lambda|}{n} t^{\frac{|\lambda|}{n} - 1} f(x_1, \ldots, x_n).$$

En faisant $t = 1$, on obtient (11).

2) Supposons maintenant que la relation (11) a lieu. Fixons un point $x = (x_1, \ldots, x_n)$ et formons la fonction

$$\varphi(t) = t^{-m\frac{|\lambda|}{n}} f(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n). \tag{12}$$

En dérivant cette fonction par rapport à $t$, on trouve

$$\varphi'(t) = \frac{t^{m\frac{|\lambda|}{n}} \sum\limits_{i=1}^{n} \frac{\partial f(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n)}{\partial x_i} \lambda_i x_i t^{\lambda_i - 1}}{t^{2m\frac{|\lambda|}{n}}} - \frac{m\frac{|\lambda|}{n} t^{m\frac{|\lambda|}{n} - 1} f(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n)}{t^{2m\frac{|\lambda|}{n}}} =$$

$$= \frac{\sum\limits_{i=1}^{n} \lambda_i t^{\lambda_i} x_i f'_{x_i}(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n) - m\frac{|\lambda|}{n} f(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n)}{t^{m\frac{|\lambda|}{n} + 1}} = 0.$$

La dernière égalité a lieu pour le point $(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n)$ en vertu de (11).

Donc, $\varphi'(t) = 0$ et $\varphi(t) = c$. La constante $c$ se détermine à partir de la condition $\varphi(1) = f(x_1, \ldots, x_n)$.

Donc, de (12) on déduit

$$f(t^{\lambda_1}x_1, \ldots, t^{\lambda_n}x_n) = t^{m\frac{|\lambda|}{n}} f(x_1, \ldots, x_n),$$

c'est-à-dire que $f$ est $\lambda$-homogène de degré $m$.

## § 9. Différentielle d'une fonction. Différentielle d'ordre supérieur

Considérons la fonction

$$W = f(x) = f(x_1, \ldots, x_n), \tag{1}$$

définie sur un ouvert $G \subset R_n$. On peut la représenter d'une infinité de manières sous la forme

$$W = \varphi(u) = \varphi(u_1, \ldots, u_m), \tag{2}$$

où

$$u_j = \psi_j(x) \quad (j = 1, \ldots, m; \ x \in G). \tag{3}$$

On se servira plus bas de la terminologie suivante: *la variable W est une fonction de la variable vectorielle indépendante $x$; la variable W est aussi fonction de la variable vectorielle dépendante $u$.* Cette dernière dépend de la variable indépendante $x$: à tout vecteur $x \in G$ correspond un vecteur $u = (\psi_1(x), \ldots, \psi_m(x))$.

Ainsi, le rôle de la variable vectorielle $x$ est exceptionnel: dans les raisonnements ci-après $x$ est traitée *uniquement comme une variable indépendante.*

Supposons qu'une fonction $f$ possède des dérivées partielles du premier ordre continues en un point $x \in G$. Nous savons qu'elle est différentiable (cf. § 6), c'est-à-dire que son accroissement en ce point peut se mettre sous la forme

$$\Delta W = \sum_{j=1}^{n} \frac{\partial W}{\partial x_j} \Delta x_j + o(\rho) \quad (\rho \to 0), \tag{4}$$

$$\rho = |\Delta x| = \Big(\sum_{j=1}^{n} \Delta x_j^2\Big)^{1/2},$$

et sa différentielle est égale à

$$dW = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \Delta x_j. \tag{5}$$

Pour les variables indépendantes $x_1, \ldots, x_n$ on admet que

$$\Delta x_j = dx_j \quad (j = 1, \ldots, n). \tag{6}$$

Ces quantités sont appelées non seulement accroissements mais aussi *différentielles* des variables indépendantes $x_j$. On les appellera *différentielles indépendantes* pour exprimer qu'elles ne dépendent pas de $x = (x_1, \ldots, x_n)$. L'« indépendance » des $dx_j$ se manifestera dans le fait que ces quantités seront traitées comme des *constantes* lors d'une différentiation par rapport à $x_1, \ldots, x_n$ $(d(dx_j) = 0)$.

En vertu de (6), on peut mettre la différentielle de $W$ sous la forme

$$dW = \sum_{j=1}^{n} \frac{\partial W}{\partial x_j} dx_j. \tag{7}$$

Il est évident que $dW$ dépend de $x_1, \ldots, x_n$ et de $dx_1, \ldots, dx_n$.

Si $u$ et $v$ sont des fonctions arbitraires possédant des dérivées partielles continues en un point $x$, alors on a

$$d(u \pm v) = du \pm dv, \tag{8}$$

$$d(uv) = u\,dv + v\,du, \tag{9}$$

$$d\left(\frac{u}{v}\right) = \frac{v\,du - u\,dv}{v^2} \quad (v \neq 0), \tag{10}$$

les dérivées partielles des fonctions entre parenthèses étant continues en $x$.

Prouvons par exemple l'égalité (10). On a

$$d\left(\frac{u}{v}\right) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{u}{v}\right) dx_j = \sum_{j=1}^{n} \frac{v\frac{\partial u}{\partial x_j} - u\frac{\partial v}{\partial x_j}}{v^2} dx_j =$$

$$= \frac{1}{v^2}\left(v \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} dx_j - u \sum_{j=1}^{n} \frac{\partial v}{\partial x_j} dx_j\right) = \frac{v\,du - u\,dv}{v^2}.$$

La continuité de $\frac{\partial}{\partial x_j}\left(\frac{u}{v}\right)$ résulte de la deuxième égalité.

La différentielle de la fonction $W$ s'appelle encore *différentielle du premier ordre*, car il existe des différentielles d'ordre supérieur.

Supposons maintenant que la fonction $W$ possède des dérivées partielles secondes continues. Par définition, la *différentielle seconde* de $W$, correspondant aux (différentielles) $dx_1, \ldots, dx_n$, est

$$d^2W = d\,(dW). \tag{11}$$

Dans cette égalité on admet que les deux opérations $d$ du second membre sont prises pour les différentielles indépendantes $dx_1, \ldots, dx_n$ considérées comme des constantes (ne dépendant pas de $x_1, \ldots, x_n$). Donc

$$d^2W = d \sum_{j=1}^{n} \frac{\partial W}{\partial x_j} dx_j = \sum_{j=1}^{n} d\left(\frac{\partial W}{\partial x_j} dx_j\right) = \sum_{j=1}^{n} \left(d\,\frac{\partial W}{\partial x_j}\right) dx_j =$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 W}{\partial x_i \partial x_j} dx_i\,dx_j. \tag{12}$$

Comme $\frac{\partial^2 W}{\partial x_i \partial x_j} = \frac{\partial^2 W}{\partial x_j \partial x_i}$, la différentielle seconde est une forme quadratique en $dx_1, \ldots, dx_n$. On appelle *forme quadratique* des variables $\xi_1, \ldots, \xi_n$ une fonction de la forme $\sum_{k=1}^{n} \sum_{i=1}^{n} a_{ik} \xi_i \xi_k$, où $a_{ik} = a_{ki}$.

D'une façon générale, la différentielle d'ordre $l$ de $W$, correspondant aux différentielles indépendantes $dx_1, \ldots, dx_n$, se définit à l'aide de la relation récurrentielle

$$d^l W = d(d^{l-1} W) \quad (l = 2, 3, \ldots), \tag{13}$$

où $d^l$, $d$ et $d^{l-1}$ sont prises pour les différentielles indiquées traitées comme des constantes (ne dépendant pas de $x_1, \ldots, x_n$).

En procédant comme dans (12), on trouve sans peine que

$$d^3 W = \sum_{k=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^3 W}{\partial x_k \partial x_i \partial x_j} dx_k \, dx_i \, dx_j.$$

L'existence des dérivées partielles continues de $f$ nous permet de simplifier l'écriture des différentielles.

Par exemple, pour la fonction de deux variables $u = f(x, y)$, on a

$$d^2 u = f''_{x^2} dx^2 + 2 f''_{xy} dx \, dy + f''_{y^2} dy^2,$$

$$d^3 u = d(d^2 u) = \frac{\partial^3 f}{\partial x^3} dx^3 + 3 \frac{\partial^3 f}{\partial x^2 \partial y} dx^2 \, dy + 3 \frac{\partial^3 f}{\partial x \, \partial y^2} dx \, dy^2 + \frac{\partial^3 f}{\partial y^3} dy^3.$$

Un raisonnement par récurrence nous donne aussitôt

$$d^n u = \frac{\partial^n f}{\partial x^n} dx^n + n \frac{\partial^n f}{\partial x^{n-1} \partial y} dx^{n-1} dy + \ldots$$
$$\ldots + \frac{n(n-1) \ldots (n-k+1)}{k!} \frac{\partial^n f}{\partial x^{n-k} \partial y^k} dx^{n-k} dy^k + \ldots + \frac{\partial^n f}{\partial y^n} dy^n,$$

ce qu'on note symboliquement

$$d^n u = d^n f = \left\{ \frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right\}^n f,$$

où il faut d'abord élever à la puissance $n$ et écrire ensuite $f$ à droite de chaque symbole $\partial^n$.

Si $f$ dépend de $m$ variables, on a une formule analogue:

$$d^m f = \left\{ \sum_{i=1}^{n} \frac{\partial}{\partial x_i} dx_i \right\}^m f. \tag{14}$$

Nous avons défini la notion de différentielle de la fonction $W$ en termes de variables indépendantes $x_1, \ldots, x_n$ (ou de la variable

vectorielle indépendante $x$). Mais traitons, ainsi que nous l'avons mentionné au début du paragraphe, $W$ comme une fonction d'une variable vectorielle *dépendante* $u = (u_1, \ldots, u_m)$. Comment dans ce cas s'expriment les différentielles première et d'ordre supérieur. Commençons par la différentielle première.

On admettra que les fonctions $\varphi(u)$ et $\psi_j(x)$ $(j = 1, \ldots, m)$ citées au début du paragraphe possèdent des dérivées partielles continues. Alors

$$dW = \sum_{j=1}^{n} \frac{\partial W}{\partial x_j} dx_j = \sum_{j=1}^{n} \left( \sum_{i=1}^{m} \frac{\partial W}{\partial u_i} \frac{\partial u_i}{\partial x_j} \right) dx_j =$$

$$= \sum_{i=1}^{m} \frac{\partial W}{\partial u_i} \sum_{j=1}^{n} \frac{\partial u_i}{\partial x_j} dx_j = \sum_{i=1}^{m} \frac{\partial W}{\partial u_i} du_i. \quad (15)$$

Donc, comme dans le cas d'une variable, la forme de la différentielle première de $W$ est la même pour les variables dépendantes et les variables indépendantes. On exprime ceci en disant que *la forme de la différentielle première est intrinsèque.*

Pour étudier ce problème pour la différentielle seconde, on admettra que les fonctions $\varphi$ et $\psi_j$ possèdent des dérivées partielles secondes continues.

En différentiant les deux membres de (15) et en tenant compte des propriétés (8) et (9), on obtient

$$d^2W = d(dW) = \sum_{i=1}^{m} d\left( \frac{\partial W}{\partial u_i} du_i \right) =$$

$$= \sum_{i=1}^{m} \left( \sum_{j=1}^{m} \frac{\partial^2 W}{\partial u_j \partial u_i} du_j \, du_i + \frac{\partial W}{\partial u_i} d^2u_i \right) =$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 W}{\partial u_j \partial u_i} du_j \, du_i + \sum_{i=1}^{m} \frac{\partial W}{\partial u_i} d^2u_i. \quad (16)$$

Dans la deuxième égalité, on s'est servi de la propriété (8), dans la troisième, de la propriété (9) et du fait aussi que la forme de la différentielle première était intrinsèque. Nous constatons que la différentielle seconde de $W$ se compose de deux termes. Le premier est une forme quadratique identique à la forme (12), où $d^2W$ est exprimé en fonction des variables indépendantes. Le second terme est une contribution dont il faut tenir compte : si $u_j$ $(j = 1, \ldots, m)$ n'est pas une fonction linéaire de $x_j$, ce terme n'est pas nul.

A noter que de ce qui vient d'être dit, il résulte que si l'expression (16) est prise pour les $dx_1, \ldots, dx_n$ de l'expression (12), alors

ces expressions sont identiquement égales quel que soit $x$ pour lequel existent les dérivées partielles continues du second ordre indiquées, et quelles que soient les $dx_i$ indépendantes.

Le calcul des différentielles $d^3W$, $d^4W$, . . . en fonction des variables dépendantes $u_j$ s'effectue de proche en proche de façon analogue. A noter que plus l'ordre est élevé, plus les expressions des différentielles seront compliquées.

## § 10. Formule de Taylor

Limitons-nous à une fonction de deux variables. Supposons qu'une fonction $u = f(x, y)$ possède au voisinage d'un point $P_0 = (x_0, y_0)$ des dérivées partielles continues jusqu'à l'ordre $l$ compris. Prenons un point $P_1 = (x_0 + \Delta x, y_0 + \Delta y)$ dans ce voisinage et joignons $P_0$ et $P_1$ par un segment de droite dont l'équation paramétrique est:

$$x = x_0 + t\,\Delta x,\ y = y_0 + t\,\Delta y\ (0 \leqslant t \leqslant 1).$$

Le long de ce segment, la fonction $u = f(x, y)$ sera fonction d'une seule variable $t$:

$$f(x, y) = f[x_0 + t\,\Delta x, y_0 + t\,\Delta y] = F(t). \tag{1}$$

Il est immédiat de voir que

$$\Delta f(P_0) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = F(1) - F(0). \tag{2}$$

La formule de Maclaurin de la fonction $F(t)$ en $t_0 = 0$ s'écrit:

$$F(t) = F(0) + \frac{F'(0)}{1!}t + \ldots + \frac{F^{(l-1)}(0)}{(l-1)!}t^{l-1} + \frac{F^{(l)}(\theta)}{l!}t^l$$
$$(0 < \theta < t).$$

En faisant $t = 1$, on obtient

$$F(1) - F(0) = \sum_{k=1}^{l-1} \frac{F^{(k)}(0)}{k!} + \frac{F^{(l)}(\theta)}{l!},\ \text{où}\ 0 < \theta < 1. \tag{3}$$

Calculons les dérivées de la fonction $F(t)$ au moyen de $f(x, y)$. De (1) on déduit

$$F'(t) = \frac{\partial f(x_0 + t\,\Delta x,\ y_0 + t\,\Delta y)}{\partial x}\Delta x + \frac{\partial f(x_0 + t\,\Delta x,\ y_0 + t\,\Delta y)}{\partial y}\Delta y,$$

d'où pour $t = 0$

$$F'(0) = \frac{\partial f(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f(x_0, y_0)}{\partial y}\Delta y = df(P_0).$$

De façon analogue,

$$F''(t) = f''_{x^2}(x_0 + t\,\Delta x,\ y_0 + t\,\Delta y)\,\Delta x^2 + 2f''_{xy}(x_0 + t\,\Delta x,\ y_0 + t\,\Delta y)\,\Delta x\,\Delta y + f''_{y^2}(x_0 + t\,\Delta x,\ y_0 + t\,\Delta y)\,\Delta y^2,\quad F''(0) = d^2 f(P_0).$$

En poursuivant ce processus on obtient

$$F'''(0) = d^3 f(P_0),\ \ldots,\ F^{(l-1)}(0) = d^{l-1} f(P_0).$$

De (2) et (3) on déduit en définitive

$$\Delta f(P_0) = \frac{df(P_0)}{1!} + \ldots + \frac{d^{l-1}f(P_0)}{(l-1)!} + \frac{1}{l!}\,d^l f(x_0 + \theta\,\Delta x,\ y_0 + \theta\,\Delta y). \quad (4)$$

La formule (4) s'appelle *formule de Taylor* pour la fonction $u = f(x, y)$. Elle ressemble à la formule de Taylor pour une fonction d'une variable, mais, développée, elle est bien plus compliquée.

La formule de Taylor d'une fonction de $n$ variables ($n > 2$) est de la même forme (4).

Pour $l = 1$, elle s'écrit

$$f(\boldsymbol{x}) - f(\boldsymbol{x}^0) = \sum_{j=1}^{n} \left(\frac{\partial f}{\partial x_j}\right)_{\boldsymbol{x}^0 + \theta(\boldsymbol{x} - \boldsymbol{x}^0)} (x_j - x_j^0) \quad (0 < \theta < 1),$$

où le symbole $(\ )_{\boldsymbol{a}}$ signifie que la fonction entre parenthèses est calculée au point $\boldsymbol{x} = \boldsymbol{a}$. Cette formule est une généralisation de celle des accroissements finis.

Pour $l = 2$ et $n = 2$, la formule (4) se détaille comme suit:

$$f(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + \frac{1}{2}\Big[\frac{\partial^2 f}{\partial x^2}(x_0 + \theta(x - x_0),\ y_0 + \theta(y - y_0))(x - x_0)^2 + 2\frac{\partial^2 f}{\partial x\,\partial y}(x_0 + \theta(x - x_0),\ y_0 + \theta(y - y_0))(x - x_0)(y - y_0) + \frac{\partial^2 f}{\partial y^2}(x_0 + \theta(x - x_0),\ y_0 + \theta(y - y_0))(y - y_0)^2\Big].$$

Pour $l = 2$ et $n$ arbitraire, la formule (4) s'écrit:

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^0) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_{\boldsymbol{x}^0} (x_i - x_i^0) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{\boldsymbol{x}^0 + \theta(\boldsymbol{x} - \boldsymbol{x}^0)} (x_i - x_i^0)(x_j - x_j^0),$$

où $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{x}^0 = (x_1^0, \ldots, x_n^0)$.

## § 11. Ensemble fermé

On dit qu'un ensemble $A \subset R_n$ est *borné* s'il existe un nombre $M > 0$ tel que

$$|x| \leqslant M, \quad \forall x \in A,$$

autrement dit s'il existe une boule de $R_n$ centrée en 0 et contenant $A$.

On dit qu'un ensemble $A$ est *fermé* si la limite de toute suite convergente de points $x^k$ ($k = 1, 2, \ldots$) de $A$ appartient à $A$.

Cette définition n'affirme pas l'existence d'une suite convergente dans $A$. Il dit simplement que, si $A$ contient une telle suite, alors sa limite appartient à $A$.

Ceci montre que l'ensemble vide est fermé. L'espace $R_n$ est visiblement fermé, mais pas borné.

Considérons à titre d'exemple l'ellipsoïde

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (a, b, c > 0), \tag{1}$$

c'est-à-dire l'ensemble des points $(x, y, z)$ vérifiant l'équation (1). Désignons cet ensemble par $B$. Cet ensemble est borné, car pour l'un quelconque de ses points on a

$$x^2 + y^2 + z^2 \leqslant m\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) = m \cdot 1 = m,$$

où $m \geqslant a^2, b^2, c^2$. Il est fermé, car si l'on considère une suite quelconque de points $(x_k, y_k, z_k) \in B$ tendant vers $(x_0, y_0, z_0)$, alors ce point appartient aussi à $B$. En effet, de

$$\frac{x_k^2}{a^2} + \frac{y_k^2}{b^2} + \frac{z_k^2}{c^2} = 1 \quad (k = 1, 2, \ldots)$$

il s'ensuit par passage à la limite pour $k \to \infty$ que

$$\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} + \frac{z_0^2}{c^2} = 1,$$

c'est-à-dire que $(x_0, y_0, z_0) \in B$.

Considérons maintenant l'ensemble des points dont les coordonnées vérifient l'inégalité

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1. \tag{2}$$

L'ensemble $A$ est aussi borné. Il est fermé, car si

$$(x_k, y_k, z_k) \in A \quad (k = 1, 2, \ldots),$$

c'est-à-dire que

$$\frac{x_k^2}{a^2}+\frac{y_k^2}{b^2}+\frac{z_k^2}{c^2}\leqslant 1,$$

et $(x_k, y_k, z_k)\to(x_0, y_0, z_0)$, alors il est évident que

$$\lim_{k\to\infty}\left(\frac{x_k^2}{a^2}+\frac{y_k^2}{b^2}+\frac{z_k^2}{c^2}\right)=\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}+\frac{z_0^2}{c^2}\leqslant 1,$$

c'est-à-dire que $(x_0, y_0, z_0)\in A$.

Il est intéressant, à ce propos, de considérer un troisième exemple d'ensemble $A'$ de points $(x, y, z)$ vérifiant l'inégalité stricte

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}<1. \tag{3}$$

L'ensemble $A'$ est ouvert (cf. § 2) mais pas fermé. Considérons par exemple une suite strictement croissante de points $(\alpha_k, 0, 0)$, où $\alpha_k$ tendent vers un nombre $a$. Alors $(\alpha_k, 0, 0)\in A'$ $(k=1, 2, \ldots)$ et $(\alpha_k, 0, 0)\to(a, 0, 0)$. Mais le point limite $(a, 0, 0)$ n'appartient pas à $A'$.

Les exemples considérés se généralisent sans peine. Soit donnée une fonction continue $F(\boldsymbol{x})=F(x_1, \ldots, x_n)$ sur l'espace $R_n$ tout entier. Alors l'ensemble $B$ de tous les points $\boldsymbol{x}=(x_1, \ldots, x_n)$ tels que

$$F(\boldsymbol{x})=F(x_1, \ldots, x_n)=C, \tag{4}$$

où $C$ est un nombre arbitraire, est fermé.

En effet, si aucun point ne vérifie (4), l'ensemble $B$ est vide, donc fermé. Supposons maintenant que $B$ n'est pas vide et qu'une suite de points $\boldsymbol{x}^k\in B$ converge vers un point $\boldsymbol{x}^0\in R_n$ (si $B$ est composé d'un seul point $\boldsymbol{x}^0$, on peut construire une suite convergente de points de $B$, plus exactement la suite $\{\boldsymbol{x}^0, \boldsymbol{x}^0, \ldots\}$). Alors $F(\boldsymbol{x}^k)=C$ $(k=1, 2, \ldots)$ et, puisque $F$ est continue en $\boldsymbol{x}^0$,

$$\lim_{k\to\infty} F(\boldsymbol{x}^k)=F(\boldsymbol{x}^0)=C.$$

Donc $\boldsymbol{x}^0\in B$ et l'ensemble $B$ est fermé.

De façon analogue, l'ensemble des points $\boldsymbol{x}$ vérifiant l'inégalité $F(\boldsymbol{x})\leqslant C$, où $C$ est un nombre quelconque et $F$ une fonction continue sur $R_n$, est fermé, car les relations

$$F(\boldsymbol{x}^k)\leqslant C \quad (k=1, 2, \ldots), \quad \boldsymbol{x}^k\to\boldsymbol{x}^0,$$

impliquent, puisque $F$ est continue sur $R_n$, que $F(\boldsymbol{x}^0)\leqslant C$.

D'après ce qui précède, l'ellipsoïde à $n$ dimensions

$$\sum_{k=1}^{n}\frac{x_k^2}{a_k^2}=1 \quad (a_k>0) \tag{5}$$

est un fermé de $R_n$.

L'ellipsoïde

$$\sum_{k=1}^{n} \frac{x_k^2}{a_k^2} \leqslant 1 \tag{6}$$

est également fermé. Par contre, l'ensemble

$$\sum_{k=1}^{n} \frac{x_k^2}{a_k^2} < 1 \tag{7}$$

n'est pas fermé. On peut s'en assurer en raisonnant comme dans le cas de la formule (3). Cet ensemble est un ellipsoïde ouvert à $n$ dimensions (cf. § 2).

Soient $A$ un ensemble quelconque de $R_n$, $x^0$ un point arbitraire de $R_n$ ($A \subset R_n$, $x^0 \in R_n$). Trois cas exclusifs peuvent se présenter:

1. Il existe une boule ouverte $V(x^0)$ de centre $x^0$, entièrement contenue dans $A$ ($V(x^0) \subset A$). Dans ce cas, $x^0$ est par définition un *point intérieur à l'ensemble* $A$ (cf. § 2).

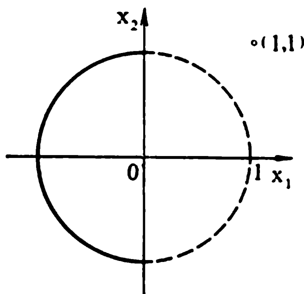



Fig. 99

2. Il existe une boule $V(x^0)$ dont aucun point n'appartient à $A$ ($V(x^0) \subset R_n \setminus A$). Alors $x^0$ est par définition un *point extérieur à l'ensemble* $A$.

3. Toute boule $V(x^0)$ contient des points de $A$ et des points n'appartenant pas à $A$. Dans ce cas, $x^0$ est par définition un *point frontière de* $A$.

L'ensemble de tous les points intérieurs à $A$ s'appelle *intérieur de* $A$ et se note Int $(A)$. C'est un ensemble ouvert (cf. § 2). Si Int $(A)$ n'est pas vide, chaque point de Int $(A)$ peut être recouvert par une boule centrée en ce point et entièrement contenue dans $A$. Si Int $(A)$ est un ensemble vide, on convient qu'il est ouvert.

L'ensemble de tous les points frontières de $A$ s'appelle *frontière de* $A$ et se note Fr $(A)$. C'est un ensemble fermé, car si $x^k \to x^0$ et $x^k \in \mathrm{Fr}(A)$ $(k = 1, 2, \ldots)$, alors toute boule ouverte $V(x^0)$ contient un point $x^k$. Ce point peut être recouvert par une boule $V(x^k)$ entièrement contenue dans $V(x^0)$ ($V(x^k) \subset V(x^0)$). Or, $V(x^k)$ contient des points de $A$ et des points n'appartenant pas à $A$, donc il en est de même de $V(x^0)$, et $x^0$ est un point frontière: $x^0 \in \mathrm{Fr}(A)$.

L'ensemble Ext $(A)$ de tous les points extérieurs à $A$ est visiblement ouvert.

Les points frontières de $A$ peuvent appartenir ou non à $A$.

L'ensemble $A \subset R_2$ de la figure 99 est composé des points $(x_1, x_2)$ tels que

$$x_1^2 + x_2^2 \leqslant 1,\ x_1 \leqslant 0;\ x_1^2 + x_2^2 < 1,\ x_1 > 0;\ x_1 = x_2 = 1.$$

Il est clair que Int $(A)$ est l'intérieur d'un disque de rayon 1 et de centre $O$; l'ensemble Fr $(A)$ est composé des points du cercle $x_1^2 + x_2^2 = 1$ et du point (1, 1); Ext $(A)$ est l'ensemble des points extérieurs au cercle unitaire, privé du point (1, 1). La moitié droite du cercle n'est pas contenue dans $A$, mais est une partie de la frontière de $A$. Signalons que l'ensemble $A$ n'est ni ouvert ni fermé.

Si donc un ensemble $A \subset R_n$ est donné, les parties Int $(A)$, Ext $(A)$ et Fr $(A)$ constituent une partition de $R_n$:

$$R_n = \text{Int}(A) + \text{Fr}(A) + \text{Ext}(A).$$

Si pour ensemble $A$ on prend l'ellipsoïde fermé à $n$ dimensions (6), alors Int $(A)$ est l'ellipsoïde ouvert (7), et Fr $(A)$, l'ellipsoïde (5).

*Si $A$ est un ouvert, alors $R_n \setminus A$ est un fermé, et réciproquement.* En effet, soit $A$ un ouvert. Supposons que $x^k \to x^0$, $x^k \in R_n \setminus A$. Si $x^0$ appartenait à $A$, alors, $A$ étant ouvert, il existerait une boule $V(x^0)$ entièrement contenue dans $A$. Or, cela est impossible, car $V(x^0)$ contient des points $x^k$ appartenant à $R_n \setminus A$. Donc, $x^0 \in R_n \setminus A$ et $R_n \setminus A$ est fermé.

Supposons maintenant que $A$ est fermé et soit $x^0 \in R_n \setminus A$. Si le point $x^0$ était un point frontière de $A$, alors toute boule $V(x^0)$ contiendrait des points de $A$. On pourrait alors construire une suite de points $x^k \in A$ convergeant vers $x^0$. L'ensemble $A$ étant fermé, $x^0$ appartiendrait à $A$, ce qui contredit l'hypothèse $x^0 \in R_n \setminus A$. Nous avons prouvé que tout point $x^0 \in R_n \setminus A$ est un point intérieur de $R_n \setminus A$, c'est-à-dire que $R_n \setminus A$ est un ouvert.

L'ensemble $A + \text{Fr}(A)$ s'appelle *adhérence* de $A$ et se note

$$\overline{A} = A + \text{Fr}(A).$$

Il est évident que

$$\text{Int}(A) + \text{Fr}(A) = A + \text{Fr}(A),$$

car d'une part, Int $(A) \subset A$ et par suite Int $(A)$ + Fr $(A) \subset A$ + Fr $(A)$, et de l'autre, si $x \in A + \text{Fr}(A)$, alors ou bien $x \in \text{Fr}(A)$ et alors $x \in \text{Int}(A) + \text{Fr}(A)$, ou bien $x \in A$ et $x \notin \text{Fr}(A)$, et alors $x \in \text{Int}(A) \subset \text{Int}(A) + \text{Fr}(A)$.

Par ailleurs, $\overline{A} = A + \text{Fr}(A)$ est un fermé, car l'extérieur de $A + \text{Fr}(A) = \text{Int}(A) + \text{Fr}(A)$ est un ouvert.

Donc, pour obtenir $\overline{A}$ il faut ajouter à $A$ tous ses points frontières ne lui appartenant pas.

Si $A$ est fermé, alors

$$A = A + \text{Fr}(A) = \overline{A},$$

autrement dit, $A$ contient tous ses points frontières. En effet, $R_n \setminus A$ est un ouvert et chaque point $x^0 \in R_n \setminus A$ peut être recouvert par une boule $V(x^0)$ ne contenant aucun point de $A$. Inversement, si

$$A = A + \text{Fr}(A) = \overline{A},$$

alors $A$ est fermé. En effet, si $x^k \to x^0$, $x^k \in A$, et si l'on admet que $x^0 \notin A$, on obtient une contradiction, car à cette condition on aurait $x^0 \in \text{Fr}(A) \subset A + \text{Fr}(A) = A$.

En conclusion, *pour qu'un ensemble $A$ soit fermé, il est nécessaire et suffisant qu'il soit confondu avec son adhérence* $(A = \overline{A})$.

En particulier, $\bar{A}$ est toujours fermée et par suite $\bar{\bar{A}} = \bar{A}$.

Notons enfin que l'ensemble vide et l'ensemble $R_n$ sont à la fois ouverts et fermés. On démontre dans les autres cas que si un ensemble est ouvert, il ne peut être fermé, et réciproquement.

## § 12. Fonction continue sur un fermé borné

Supposons provisoirement que $A$ est un ensemble dans un espace $R_n$ et soit $f(x) = f(x_1, \ldots, x_n)$ une fonction définie sur $A$. *La fonction $f$ est par définition continue en un point $x^0 \in A$ si*

$$\lim_{\substack{x^k \to x^0 \\ x^k \in A}} f(x^k) = f(x^0) \tag{1}$$

*pour toute suite de points $x^k \in A$ convergeant vers $x^0$.*

Dans la définition de la continuité, donnée au § 4, on a admis que la fonction $f$ est définie dans un voisinage du point $x^0$ et l'on a exigé que la limite (1) existe et soit égale à $f(x^0)$ pour *toute suite de points $x^k$ convergeant vers $x^0$.*

Ici on n'exige pas que la fonction $f$ soit définie dans un voisinage du point $x^0$, mais simplement qu'elle soit définie au point $x^0 \in A$ et que la limite (1) ait lieu pour toute suite de points $x^k \in A$ convergeant vers $x^0$.

La définition ci-dessus peut être formulée en termes de $\varepsilon$ et $\delta$: la fonction $f$ est continue en un point $x^0 \in A$ si pour tout $\varepsilon > 0$ on peut exhiber un $\delta > 0$ tel que

$$|f(x) - f(x^0)| < \varepsilon, \quad \forall x \in A, \quad |x - x^0| < \delta.$$

Supposons maintenant que $A$ est un ensemble fermé borné dans $R_n$ et que $f(x)$ est une fonction définie et continue sur $A$. Sous ces conditions on peut démontrer les propriétés remarquables suivantes:

1) *La fonction $f$ est bornée sur $A$.*

2) *La fonction $f$ passe par son maximum et son minimum sur $A$, autrement dit il existe dans $A$ des points $x^0$ et $y^0$ tels que*

$$f(x^0) = \max_{x \in A} f(x), \quad f(y^0) = \min_{x \in A} f(x).$$

3) *La fonction $f$ est uniformément continue sur $A$, c'est-à-dire que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que*

$$|f(x') - f(x'')| < \varepsilon$$

*quels que soient $x', x'' \in A$ tels que $|x' - x''| < \delta$.*

On remarquera que les propriétés 1), 2) et 3) généralisent les propriétés notoires d'une fonction continue $f(x)$ définie sur $[a, b]$. Signalons que l'intervalle $[a, b]$ est un ensemble fermé borné à une

dimension. En effet, si une suite quelconque de points $x_k \in [a, b]$ converge vers un point $x_0$, alors celui-ci appartient aussi à $[a, b]$.

La démonstration des propriétés 1), 2) et 3) est exactement la même que celle exhibée aux §§ 5 et 7 du chap. 3 pour l'intervalle $[a, b]$. Elle repose entièrement sur le lemme suivant qui généralise le théorème de Bolzano-Weierstrass du § 9, chap. 2.

LEMME. *De toute suite bornée de points* $x^k = (x_1^k, \ldots, x_n^k)$ $(k = 1, 2, \ldots)$ *on peut extraire une suite* $\{x^{k_l}\}$ $(l = 1, 2, \ldots)$ *convergeant vers un point* $x^0$:

$$|x^{k_l} - x^0| \to 0 \quad (l \to \infty).$$

DÉMONSTRATION. La suite $\{x^k\}$ étant bornée, il existe un nombre $M$ tel que

$$M \geqslant |x^k| \geqslant |x_j^k| \quad (j = 1, \ldots, n;\ k = 1, 2, \ldots).$$

Ceci exprime que les coordonnées des points $x^k$ sont bornées, elles aussi. Les premières coordonnées forment une suite bornée $\{x_1^k\}$ $(k = 1, 2, \ldots)$. En vertu du théorème de Bolzano-Weierstrass on peut en extraire une suite $\{x_1^{k_{l_1}}\}$ convergeant vers un nombre $x_1^0$. Considérons les deuxièmes coordonnées $x_2^k$ uniquement pour les indices $k_{l_1}$. La suite extraite $\{x_2^{k_{l_1}}\}$ étant bornée, on peut en extraire une suite $\{x_2^{k_{l_2}}\}$ convergeant vers un nombre $x_2^0$. Comme $\{k_{l_2}\}$ est une suite extraite de $\{k_{l_1}\}$, on a simultanément $x_1^{k_{l_2}} \to x_1^0$, $x_2^{k_{l_2}} \to x_2^0$. En poursuivant ce processus, on obtient une suite extraite $k_{l_n} = k_l$ et un ensemble de nombres $x_1^0, x_2^0, \ldots, x_n^0$ tels que l'on ait simultanément

$$x_1^{k_l} \to x_1^0,\ x_2^{k_l} \to x_2^0,\ \ldots,\ x_n^{k_l} \to x_n^0 \quad (l \to \infty).$$

En posant $x^0 = (x_1^0, \ldots, x_n^0)$ on obtient ce qu'on voulait.

DÉMONSTRATION DE 1). Supposons que $f$ n'est pas bornée sur l'ensemble fermé borné $A$. Alors pour tout entier naturel $m$ il existe un point $x^m \in A$ tel que

$$|f(x^m)| > m \quad (m = 1, 2, \ldots). \tag{2}$$

L'ensemble $A$ étant borné, il en est de même de la suite $\{x^m\}$, et en vertu du lemme précédent on peut en extraire une suite $\{x^{m_k}\}$ convergeant vers un point $x^0$. L'ensemble $A$ étant fermé par hypothèse, le point $x^0 \in A$. Or, la fonction $f$ est continue en $x^0$, donc

$$\lim_{\substack{k \to \infty \\ x^{m_k} \in A}} f(x^{m_k}) = f(x^0). \tag{3}$$

La propriété (3) contredit la propriété (2), donc $f$ ne peut être que bornée sur un ensemble fermé borné.

DÉMONSTRATION DE 2). En vertu de la propriété 1) une fonction continue sur un fermé borné $A$ est bornée, donc majorée par un

nombre $K$:

$$f(x) \leqslant K \quad (x \in A).$$

La fonction $f$ admet alors une borne supérieure sur $A$:

$$\sup_{x \in A} f(x) = M. \tag{4}$$

Le nombre $M$ est doué de la propriété suivante: pour tout entier naturel $m$ il existe un point $x^m = (x_1^m, \ldots, x_n^m) \in A$ tel que

$$M - \frac{1}{m} < f(x^m) \leqslant M \quad (m = 1, 2, \ldots).$$

La suite $\{x^m\}$ est bornée, car appartenant à l'ensemble fermé borné $A$:

$$|x^m| = \sqrt{\sum_{i=1}^{n} |x_i^m|^2} \leqslant K_1, \quad m = 1, 2, \ldots$$

On peut donc en extraire une suite $\{x^{m_k}\}$ convergeant vers un point $x^0 \in A$. Ceci résulte du fait que $A$ est fermé.

Etant continue sur $A$, la fonction $f$ l'est en $x^0$ et par suite

$$\lim_{\substack{k \to \infty \\ x^{m_k} \in A}} f(x^{m_k}) = f(x^0).$$

D'autre part,

$$M - \frac{1}{m_k} < f(x^{m_k}) \leqslant M \quad (k = 1, 2, \ldots).$$

En passant à la limite pour $k \to \infty$, on trouve

$$M \leqslant f(x^0) \leqslant M,$$

c'est-à-dire que

$$f(x^0) = M.$$

Donc, la fonction $f$ passe par son maximum au point $x^0 \in A$. Nous avons ainsi montré qu'il existait un point $x^0 \in A$ tel que

$$\max_{x \in A} f(x) = f(x^0).$$

La démonstration est identique pour le minimum.

Démonstration de 3). Si par absurde on suppose que cette propriété est fausse, alors il existerait un $\varepsilon > 0$ tel que pour tout $\delta > 0$ on pourrait exhiber un couple de points $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in A$ vérifiant l'inégalité

$$|x - y| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} < \delta,$$

et pour lesquels

$$| f(x) - f(y) | \geqslant \varepsilon.$$

Si l'on considère maintenant une suite d'entiers strictement positifs $\delta_m \to 0$, $m \to \infty$, alors pour tout $\delta_m$ il existerait des points $x^m = (x_1^m, \ldots, x_n^m)$, et $y^m = (y_1^m, \ldots, y_n^m) \in A$ tels que

$$| x^m - y^m | < \delta_m,$$

mais

$$| f(x^m) - f(y^m) | \geqslant \varepsilon. \tag{5}$$

La suite $\{x^m\}$ est bornée, car appartenant à l'ensemble borné $A$. On peut donc en extraire une suite $\{x^{m_k}\}$ convergeant vers un point $x^0 \in A$.

Comme $| x^{m_k} - y^{m_k} | \to 0$ pour $k \to \infty$, la suite extraite $\{y^{m_k}\}$ converge aussi vers $x^0$:

$$| y^{m_k} - x^0 | = | y^{m_k} - x^{m_k} + x^{m_k} - x^0 | \leqslant \leqslant | y^{m_k} - x^{m_k} | + | x^{m_k} - x^0 |.$$

La fonction $f$ est continue sur $A$ par hypothèse, donc elle l'est en $x^0$.

Signalons que la continuité de $f$ est comprise au sens de (1). Donc

$$\lim_{\substack{k \to \infty \\ x^{m_k} \in A}} f(x^{m_k}) = \lim_{\substack{k \to \infty \\ y^{m_k} \in A}} f(y^{m_k}) = f(x^0).$$

En passant à la limite dans (5) pour $k \to \infty$, on obtient

$$\varepsilon \leqslant \lim_{\substack{k \to \infty \\ x^{m_k},\, y^{m_k} \in A}} |f(x^{m_k}) - f(y^{m_k})| = |f(x^0) - f(x^0)| = 0,$$

i.e. $\varepsilon \leqslant 0$. Cette contradiction prouve la propriété 3).

## § 13. Extrémums

Soient donnés une fonction $u = f(x)$, $x = (x_1, \ldots, x_n)$, sur un domaine (i.e. un ensemble connexe ouvert) $G$ et un point $x^0 = (x_1^0, \ldots, x_n^0)$ de $G$. On dit que la fonction $u = f(x)$ passe par un *maximum local* (resp. *minimum local*) au point $x^0$ s'il existe un voisinage $V(x^0)$ tel que pour tout $x \in V(x^0)$ on a

$$f(x) \leqslant f(x^0) \quad (\text{resp. } f(x) \geqslant f(x^0)). \tag{1}$$

Le point $x^0$ s'appelle *point de maximum* (resp. *minimum*) *local*, et les valeurs $f(x^0)$, *maximum* (resp. *minimum*) *de* $f$. Le maximum et

le minimum locaux sont groupés sous l'appellation commune d'*extrémums locaux*. De la définition de l'extrémum il s'ensuit que l'accroissement $\Delta u = f(x) - f(x^0)$ ne change pas de signe dans un voisinage assez petit de $x^0$:

$$\Delta u \geqslant 0 \text{ dans le cas d'un minimum local;}$$

$$\Delta u \leqslant 0 \text{ dans le cas d'un maximum local.}$$

THÉORÈME (CONDITION NÉCESSAIRE D'EXTRÉMUM). *Supposons qu'une fonction $u = f(x)$ présente un extrémum local en un point $x^0$. Si les dérivées partielles premières $\frac{\partial f}{\partial x_i}$ $(i = 1, \ldots, n)$ existent en $x^0$, alors elles s'annulent en ce point:*

$$\frac{\partial f(x^0)}{\partial x_i} = 0 \quad (i = 1, \ldots, n). \tag{2}$$

DÉMONSTRATION. Prouvons que $\frac{\partial f(x^0)}{\partial x_1} = 0$. Figeons les variables $x_2 = x_2^0, \ldots, x_n = x_n^0$. On obtient alors une fonction $u = f(x_1, x_2^0, \ldots, x_n^0)$ d'une seule variable $x_1$, qui présente de plus un extrémum local en $x_1^0$. En vertu de la condition nécessaire d'extrémum pour une fonction d'une variable, la dérivée de la fonction $u$ par rapport à $x_1$ est nulle au point $x_1^0$. Or, cette dérivée est la dérivée partielle de la fonction $f(x)$ par rapport à $x_1$ au point $x^0$, c'est-à-dire que

$$\frac{\partial f(x_1^0, x_2^0, \ldots, x_n^0)}{\partial x_1} = \frac{\partial f(x^0)}{\partial x_1} = 0.$$

Les autres cas se traitent de façon analogue.

COROLLAIRE. *Si une fonction $u = f(x)$ présente un extrémum en un point $x^0$ et est différentiable en $x^0$, alors $df(x^0) = 0$ ou* grad $f(x^0) = \mathbf{0}$.

Ce corollaire résulte de la définition de la différentielle et du gradient.

REMARQUE. La condition (2) n'est pas suffisante pour que la fonction $f$ présente un extrémum en $x^0$.

Par exemple, la fonction $u = x^2y$ possède les dérivées partielles $\frac{\partial u}{\partial x} = 2xy$, $\frac{\partial u}{\partial y} = x^2$ qui s'annulent en (0, 0). Mais ce point n'est pas un point d'extrémum, car $\Delta u = x^2y - 0 = x^2y$ prend des valeurs aussi bien positives que négatives dans tout voisinage de ce point.

On appelle *points stationnaires* les points en lesquels les dérivées partielles continues de $f$ s'annulent.

Passons maintenant aux conditions suffisantes d'extrémum. Supposons qu'une fonction $u = f(x)$ possède des dérivées partielles premières et secondes (par rapport à toutes les variables) continues

et soit $x^0$ un point stationnaire ($df(x^0) = 0$). En développant la fonction $u = f(x)$ suivant la formule de Taylor au point $x^0$, on obtient

$$\Delta f(x^0) = df(x^0) + \frac{1}{2!} d^2 f(x^0 + \theta\,\Delta x) = \frac{1}{2} d^2 f(x^0 + \theta\,\Delta x) =$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f''_{x_i x_j}(x^0 + \theta\,\Delta x)\,\Delta x_i\,\Delta x_j =$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (f''_{x_i x_j}(x^0) + \varepsilon_{ij})\,\Delta x_i\,\Delta x_j =$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f''_{x_i x_j}(x^0)\,\Delta x_i\,\Delta x_j + \frac{\rho^2}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_{ij} \frac{\Delta x_i}{\rho} \cdot \frac{\Delta x_j}{\rho} =$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f''_{x_i x_j}(x^0)\,\Delta x_i\,\Delta x_j + \frac{\rho^2}{2} \alpha(\Delta x),$$

où $0 < \theta < 1$, $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$, $\rho = |\Delta x| = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}$,

$\alpha(\Delta x) \to 0$ pour $\rho \to 0$.

Les dérivées secondes étant continues, les quantités $\varepsilon_{ij}$ qui dépendent de $\Delta x$ tendent vers 0 pour $\rho = |\Delta x| \to 0$. Donc

$\max_{i,j} |\varepsilon_{ij}| = \varepsilon \to 0$ avec $\rho$. Comme $\frac{|\Delta x_i|}{\rho} \leqslant 1$, on obtient

$$|\alpha(\Delta x)| = \left| \sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_{ij} \frac{\Delta x_i}{\rho} \frac{\Delta x_j}{\rho} \right| \leqslant \varepsilon \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} 1 \cdot 1 = n^2 \varepsilon \to 0$$

$$(\rho \to 0).$$

Nous avons ainsi prouvé que

$$\Delta f(x^0) = f(x^0 + \xi) - f(x^0) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \xi_i \xi_j + \frac{\rho^2}{2} \alpha(\xi), \qquad (3)$$

où

$$a_{ij} = a_{ji} = f''_{x_i x_j}(x^0), \quad \xi_i = \Delta x_i, \quad \xi = (\xi_1, \ldots, \xi_n)$$

et

$$\alpha(\xi) \to 0 \quad \text{pour} \quad \rho = |\xi| = \sqrt{\sum_{i=1}^{n} \xi_i^2} \to 0.$$

L'expression

$$A(\xi)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\xi_i\xi_j \quad (a_{ij}=a_{ji}) \tag{4}$$

est une forme quadratique en $\xi = (\xi_1, \ldots, \xi_n)$. Son signe nous permet, grâce à (3), de connaître celui de $\Delta f(x^0)$ pour $|\Delta x|$ assez petits.

On a les assertions suivantes:

1) *Si la forme $A(\xi)$ est strictement positive, i.e. $A(\xi) > 0$ pour tous les $\xi \neq 0$, alors la fonction $f$ présente un minimum local au point $x^0$.*

2) *Si la forme $A(\xi)$ est strictement négative, i.e. $A(\xi) < 0$ pour tous les $\xi \neq 0$, alors la fonction $f$ présente un maximum local en $x^0$.*

3) *Si $A(\xi) \geqslant 0$ ou $A(\xi) \leqslant 0$ pour tous les $\xi$ et si existe un $\xi \neq 0$ pour lequel $A(\xi) = 0$, le problème reste ouvert.*

4) *Si le signe de la forme $A(\xi)$ n'est pas défini, c'est-à-dire si existent des vecteurs $\xi'$ et $\xi''$ pour lesquels $A(\xi') > 0$ et $A(\xi'') < 0$, alors la fonction $f$ ne présente pas d'extrémum local au point $x^0$.*

DÉMONSTRATION DE 1). Mettons l'égalité (3) sous la forme

$$\Delta f(x^0)=\frac{\rho^2}{2}\Big[\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\frac{\xi_i}{\rho}\frac{\xi_j}{\rho}+\alpha(\xi)\Big]=$$

$$=\frac{\rho^2}{2}\Big[\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\eta_i\eta_j+\alpha(\xi)\Big]=\frac{\rho^2}{2}[A(\eta)+\alpha(\xi)], \tag{5}$$

où nous avons fait le changement de variables

$$\eta_i=\xi_i/\rho \quad (i=1, 2, \ldots, n).$$

Il est immédiat de voir que

$$|\eta|=\sqrt{\sum_{i=1}^{n}\eta_i^2}=\frac{\sqrt{\sum_{i=1}^{n}\xi_i^2}}{\rho}=1.$$

Donc, quel que soit $\xi$, le point $\eta$ se trouve sur la sphère unité à $n$ dimensions. La fonction $A(\eta)$ est continue sur cette surface qui est un ensemble fermé borné. De plus, elle est strictement positive par hypothèse. Donc, elle passe par son minimum $m$ en un point de cette surface et $m > 0$ (cf. § 12, propriété 2)). Comme $\alpha(\xi) \to 0$ pour $\rho = |\xi| \to 0$, on a pour $\delta > 0$ assez petit

$$|\alpha(\xi)| < m, \quad \forall \xi : |\xi| < \delta.$$

Donc

$$\Delta f(x^0) = f(x^0+\xi) - f(x^0) =$$
$$= \frac{\rho^2}{2}[A(\eta)+\alpha(\xi)] \geqslant \frac{\rho^2}{2}[m+\alpha(\xi)] \geqslant 0, \quad \forall \xi : |\xi| < \delta,$$

et la fonction $f$ a un minimum local en $x^0$.

L'assertion 2) se démontre de façon analogue.

Démonstration de 3). La forme $A$ ($\xi$) s'annule pour un $\xi' \neq 0$. Comme elle est homogène ($A(\alpha\xi) = \alpha^2 A(\xi)$), elle s'annule aussi pour $\xi = \alpha\xi'$, où $\alpha$ est un nombre. Ceci montre que pour de tels $\xi$ la forme est nulle et par suite $f(x^0+\xi) - f(x^0) = \frac{1}{2}\rho^2\alpha(\xi)$. Mais le signe de $\alpha(\xi)$ est inconnu, donc on ne peut dire si $f$ présente ou non un extrémum en $x^0$.

Démonstration de 4). On a encore intérêt à se servir de l'égalité (5). Par hypothèse, il existe un point $\xi'$ tel que $A(\xi') > 0$ et un point $\xi''$ tel que $A(\xi'') < 0$. Donc, pour les points $\eta' = \xi'/\rho$ et $\eta'' = \xi''/\rho$ on aura $A(\eta') > 0$ et $A(\eta'') < 0$, et pour les $\rho$ assez petits, $A(\eta') + \alpha(\xi') > 0$, $A(\eta'') + \alpha(\xi'') < 0$, autrement dit tout voisinage assez petit de $x^0$ contient des points $x'$ et $x''$ pour lesquels $f(x') > f(x^0)$ et $f(x'') < f(x^0)$, ce qui exprime l'absence d'extrémum.

Le théorème de Sylvester [1]) nous permet de déterminer le signe de la forme quadratique (4) au moyen des coefficients $a_{ij}$. On signalera ici seulement des critères découlant de ce théorème pour le cas d'une fonction $u = f(x_1, x_2)$ de deux variables.

Si $a_{11} = f''_{x_1^2}(x^0) > 0$ et

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}^2 = f''_{x_1^2}(x^0)\, f''_{x_2^2}(x^0) - [f''_{x_1x_2}(x^0)]^2 > 0$$

(la forme (4) est définie strictement positive), alors la fonction $u = f(x_1, x_2)$ a un minimum local au point $x^0 = (x_1^0, x_2^0)$.

Si

$$f''_{x_1^2}(x^0) < 0, \quad f''_{x_1^2}(x^0)\, f''_{x_2^2}(x^0) - [f''_{x_1x_2}(x^0)]^2 > 0$$

(la forme (4) est définie strictement négative), alors la fonction $u = f(x_1, x_2)$ a un maximum local en $x^0$.

Si $a_{11}a_{22} - a_{12}^2 < 0$, alors le signe de $d^2f(x^0)$ traitée comme une forme quadratique n'est pas défini lorsque $\Delta x_i$ varie. Donc, $\Delta f(x^0)$ ne garde son signe dans aucun voisinage de $x^0$ et par suite $f$ n'a pas d'extrémum en $x^0$.

Si $a_{11}a_{22} - a_{12}^2 = 0$, le problème de l'extrémum reste entier.

---

[1]) Voir § 22 du chap. 10.

EXEMPLE 1. Soit la fonction $u = x^3 - 3x + y^2$. Les points $(\pm 1, 0)$ sont des points stationnaires. Voyons si la fonction $u$ y présente un extrémum. On a

$$u''_{x^2} = 6x,\ u''_{x^2}(\pm 1,\ 0) = \pm 6,\ u''_{xy} = 0,\ u''_{y^2} = 2.$$

Pour le point (1, 0) on a $a_{11}a_{22} - a_{12}^2 = 6 \cdot 2 - 0 = 12 > 0$, $a_{11} = 6 > 0$. Donc, au point (1, 0) la fonction $u$ présente un minimum local. Pour le point (—1, 0) on a $a_{11} = -6 < 0$, $a_{11}a_{22} - a_{12}^2 = -12 < 0$, donc la fonction $u$ n'a pas d'extrémum en (—1, 0).

EXEMPLE 2. Le point (0, 0) est un point stationnaire pour la fonction $u = x^4 + y^2$ et on voit sans peine que c'est un point de minimum local. Cependant, $u''_{x^2} = 12x^2$, $u''_{xy} = 0$, $u''_{y^2} = 2$, c'est-à-dire que $a_{11} = 0$, $a_{12} = 0$, $a_{22} = 2$ et $a_{11}a_{22} - a_{12}^2 = 0$.

EXEMPLE 3. Au point (0, 0) qui est stationnaire pour la fonction $u = x^3 + y^2$ on a $a_{11} = 0$, $a_{12} = 0$, $a_{22} = 2$, c'est-à-dire que $a_{11}a_{22} - a_{12}^2 = 0$, mais dans ce cas la fonction $u = x^3 + y^2$ n'a pas d'extrémum en (0, 0), car l'accroissement $\Delta u = x^3$ suivant la droite $y = 0$ change de signe en passant par le point $x = 0$.

## § 14. Calcul du minimum et du maximum d'une fonction

Soit donnée une fonction $u = f(x)$ continûment dérivable sur un ensemble $\overline{G} \subset R_n$, adhérence d'un domaine borné $G$, c'est-à-dire un domaine auquel on a ajouté la frontière de $G$. La fonction $f$ passe alors par son maximum et son minimum en des points $x \in \overline{G}$ (cf. § 12, propriété 2)). Ces points peuvent être intérieurs ou frontières. Si $x$ est un point intérieur, la fonction $f(x)$ y présente un extrémum local. Donc, pour trouver le maximum (resp. minimum) d'une fonction, il est nécessaire de trouver tous les points stationnaires, de calculer la valeur de cette fonction en ces points et de comparer ces valeurs avec celles prises sur la frontière de $G$. La plus grande de ces valeurs s'appelle *maximum* de $f$ sur $\overline{G}$.

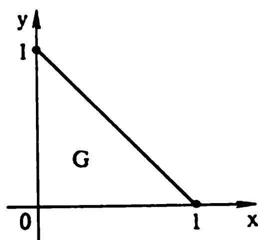


Fig. 100

Si $\overline{G} \subset R_2$ et Fr $(G)$ est une courbe continue d'équation $x = \varphi(t)$, $y = \psi(t)$, alors la fonction $f = f[\varphi(t), \psi(t)]$ est une fonction d'une variable dont on sait calculer le maximum.

EXEMPLE. Trouver le maximum de la fonction $z = 1 - x + x^2 + 2y$ définie sur l'adhérence du domaine $G$ limité par les droites: $x = 0$, $y = 0$, $x + y = 1$ (fig. 100).

SOLUTION. $z'_x = -1 + 2x = 0$, $z'_y = 2 \neq 0$, c'est-à-dire que cette fonction ne possède pas de points stationnaires. Etudions la fonction $f$ sur Fr $(G)$.

1) Soit $x = 0$. Alors $z = 1 + 2y$, $0 \leqslant y \leqslant 1$. La fonction $z = 1 + 2y$ ne possède pas de points stationnaires sur $[0, 1]$ et $z(0) = 1$, $z(1) = 3$.

2) Soit $y = 0$. Alors $z = 1 - x + x^2$, $0 \leqslant x \leqslant 1$. Par ailleurs, $z'_x = -1 + 2x = 0$ pour $x = 1/2$, c'est-à-dire que $x = 1/2$ est un point stationnaire. Calculons la valeur de la fonction en ce point et aux extrémités de $[0, 1]$. On a $z(1/2) = 3/4$, $z(0) = 1$, $z(1) = 1$.

3) Soit $x + y = 1$. Alors $z = 3 - 3x + x^2$, $0 \leqslant x \leqslant 1$. Comme $z'_x = -3 + 2x = 0$ en $x = 3/2 \notin [0, 1]$, alors il n'existe pas de points stationnaires dans $[0, 1]$. D'autre part, $z(0) = 3$, $z(1) = 1$.

En comparant les maximums de la fonction $f$ sur les diverses parties de Fr $(G)$, on constate que le maximum de la fonction $z(x, y)$ sur $\overline{G}$ est égal à 3 et est réalisé au point $x^0 = (0, 1)$.

## § 15. Théorème d'existence des fonctions implicites

Soit une fonction $f(x, y)$ de deux variables $x$ et $y$. Egalons-la à 0:

$$f(x, y) = 0. \tag{1}$$

Désignons par $\mathfrak{M}$ l'ensemble des points $(x, y)$ vérifiant (1). Soit $(x_0, y_0) \in \mathfrak{M}$, i.e. $f(x_0, y_0) = 0$.

Si $f$ n'est astreint à aucune condition, l'ensemble $\mathfrak{M}$ peut être de nature quelconque. Si par exemple $f(x, y) = (x - x_0)^2 + (y - y_0)^2$, alors l'ensemble $\mathfrak{M}$ est composé du seul point $(x_0, y_0)$. Si $f(x, y) = x^2 + y^2 + 1$, l'ensemble $\mathfrak{M}$ est vide. Si enfin

$$f(x, y) = (x - x_0)^2 - (y - y_0)^2 = \\ = (x + y - x_0 - y_0)(x - y - x_0 + y_0),$$

alors $\mathfrak{M}$ est un couple de droites passant par $(x_0, y_0)$. Cependant, il arrive souvent que $\mathfrak{M}$ est, au moins dans un voisinage assez petit de $(x_0, y_0)$, une courbe définie par une fonction continue

$$y = \psi(x), \quad x \in ]x_0 - \delta, x_0 + \delta[$$

(donc, $\psi$ est une fonction définie par (1), cf. aussi § 1 du chap. 3).

A la question comment reconnaître ce cas d'après les propriétés de la fonction répondent les deux théorèmes généraux qui suivent.

Théorème 1. *Soit donnée l'équation*

$$f(x, y) = 0. \tag{1}$$

***Supposons que la fonction $f$, définie dans un voisinage à deux dimensions $\Omega$ d'un point $(x_0, y_0)$ d'un plan rapporté à un système de coordonnées $(Ox, Oy)$, est continue dans ce voisinage avec ses dérivées***

*partielles premières, et de plus*

$$f'_y(x_0, y_0) = \left(\frac{\partial f}{\partial y}\right)_0 \neq 0, \quad f(x_0, y_0) = 0. \tag{2}$$

*Supposons d'autre part que* $\mathfrak{M}$ *est l'ensemble de tous les points* $(x, y)$ *vérifiant l'équation* (1) *(en particulier* $(x_0, y_0) \in \mathfrak{M}$*).*
*Alors, quel que soit* $b_0 > 0$, *il existe un rectangle*

$$\Delta = \{|x - x_0| < a, \ |y - y_0| < b\}, \quad b < b_0, \tag{3}$$

*contenu dans* $\Omega$ *et tel que l'ensemble* $\mathfrak{M} \cap \Delta$ *soit défini par la fonction continûment dérivable*

$$y = \psi(x), \quad x \in \Delta^0, \tag{4}$$

$$\Delta^0 = \{|x - x_0| < a\}. \tag{5}$$

En d'autres termes, le rectangle $\Delta$ est tel que sur sa projection $\Delta^0$ sur l'axe $Ox$ on peut définir une fonction continûment dérivable (4),

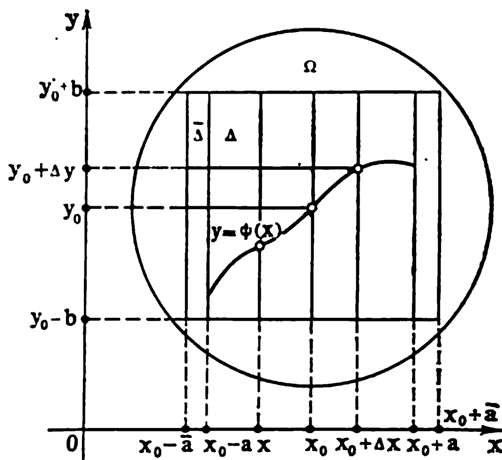



Fig. 101

solution de l'équation (1), c'est-à-dire que

$$f(x, \psi(x)) \equiv 0, \quad x \in \Delta^0. \tag{6}$$

Le graphe de cette fonction est entièrement contenu dans $\Delta$. Cette fonction est unique en ce sens que les coordonnées de tout point $(x, y) \in \mathfrak{M} \cap \Delta$ sont reliées par la relation (4). En particulier $y_0 = \psi(x_0)$, car $(x_0, y_0) \in \mathfrak{M} \cap \Delta$ (fig. 101).

DÉMONSTRATION DU THÉORÈME 1. Supposons pour fixer les idées que $f'_y(x_0, y_0) > 0$. La dérivée $f'_y$ étant continue sur $\Omega$, il existe

un voisinage du point $(x_0, y_0)$, que nous désignerons encore par $\Omega$, dans lequel $f'_y(x, y) > 0$. Considérons le rectangle fermé

$$\bar{\Delta} = \{|x - x_0| \leqslant \bar{a}, \ |y - y_0| \leqslant b\} \subset \Omega \qquad (b < b_0).$$

Alors $f'_y(x, y) > 0$ sur $\bar{\Delta}$ et

$$\min_{(x,y)\in\bar{\Delta}} f'_y(x, y) = m > 0. \tag{7}$$

La fonction $f(x, y)$ traitée comme une fonction de $y$ sur l'intervalle $[y_0 - b, y_0 + b]$, $x = x_0$, est strictement croissante et s'annule en $y = y_0$ (par hypothèse, $f(x_0, y_0) = 0$). Donc,

$$f(x_0, y_0 - b) < 0, \quad f(x_0, y_0 + b) > 0.$$

La fonction $f$ étant continue, il existe un nombre $a$ assez petit, $0 < a < \bar{a}$, tel que

$$f(x, y_0 - b) < 0, \ f(x, y_0 + b) > 0, \ \forall x \in \Delta^0 = \{|x - x_0| < a\}.$$

Considérons le rectangle ouvert $\Delta = \{|x - x_0| < a, \ |y - y_0| < < b\}$. Il est évident que $\Delta \subset \bar{\Delta} \subset \Omega$ et $\Delta^0$ est la projection de $\Delta$ sur $Ox$.

Traitons maintenant la fonction $f$ comme une fonction de $y$ sur l'intervalle $[y_0 - b, y_0 + b]$ pour $x \in \Delta^0$ arbitraire. La fonction $f$ est strictement croissante, continue ($f'_y > 0$) et possède des signes contraires aux extrémités de l'intervalle considéré. Donc, le théorème des valeurs intermédiaires affirme l'existence d'un nombre unique $y$ appartenant à l'intervalle $]y_0 - b, y_0 + b[$, que nous désignerons par $y = \psi(x)$, tel que $f(x, \psi(x)) = 0$.

Ceci prouve l'existence d'une fonction $\psi(x)$ définie sur $\Delta^0$ et vérifiant l'équation (6).

Montrons que la fonction $\psi(x)$ est continue sur $\Delta^0$. Soient $x, x + \Delta x \in \Delta^0$, $y = \psi(x)$, $\Delta y = \psi(x + \Delta x) - \psi(x)$. La formule de Taylor nous donne (cf. page 283)

$$0 = f(x + \Delta x, y + \Delta y) - f(x, y) = f'_x(x + \theta\,\Delta x, y + \theta\,\Delta y)\,\Delta x + f'_y(x + \theta\,\Delta x, y + \theta\,\Delta y)\,\Delta y,$$

où $0 < \theta < 1$. D'où

$$\frac{\Delta y}{\Delta x} = -\frac{f'_x(x+\theta\,\Delta x, \ y+\theta\,\Delta y)}{f'_y(x+\theta\,\Delta x, \ y+\theta\,\Delta y)}, \tag{8}$$

où $(x + \theta\,\Delta x, y + \theta\,\Delta y) \in \Delta$. Etant continue par hypothèse, la fonction $f'_x$ est bornée ($|f'_x| \leqslant M$) sur le rectangle fermé $\bar{\Delta}$, donc sur $\Delta \subset \bar{\Delta}$. En vertu de (7), la fonction $f'_y$ est minorée par $m > 0$, donc on déduit à partir de (8) que

$$\left|\frac{\Delta y}{\Delta x}\right| \leqslant \frac{M}{m},$$

autrement dit $\Delta y \to 0$ avec $\Delta x$, ce qui exprime que la fonction $y = \psi(x)$ est continue au point $x$. Le point $x$ étant arbitraire, la fonction $\psi(x)$ est continue sur $\Delta^0$.

En passant à la limite dans (8) pour $\Delta x \to 0$, on trouve

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\frac{f'_x(x, y)}{f'_y(x, y)} \quad (y = \psi(x)). \tag{9}$$

Nous avons prouvé l'existence de la dérivée $\psi'(x)$ en $x$ et l'égalité

$$\psi'(x) = -\frac{f'_x(x, \psi(x))}{f'_y(x, \psi(x))}. \tag{10}$$

La continuité de $\psi'(x)$ résulte immédiatement de (10), car $f'_x$ et $f'_y$ sont continues sur le rectangle $\Delta$ et la courbe $y = \psi(x)$ est entièrement contenue et continue dans $\Delta$ comme nous l'avons montré plus haut.

Formulons un théorème identique au théorème 1 dans le cas où la fonction implicite dépend de $n$ variables.

Théorème 1'. *Soit donnée l'équation*

$$f(x, y) = f(x_1, \ldots, x_n, y) = 0. \tag{1'}$$

*Supposons que la fonction $f$ est définie dans un voisinage $\Omega$ d'un point $(x^0, y^0) = (x_1^0, \ldots, x_n^0, y^0)$ de l'espace $R_{n+1}$ des points $(x, y) = (x_1, \ldots, x_n, y)$, continue avec ses dérivées partielles premières sur $\Omega$ et*

$$f'_y(x^0, y^0) = \left(\frac{\partial f}{\partial y}\right)_0 \neq 0, \quad f(x^0, y^0) = 0. \tag{2'}$$

*Supposons d'autre part que $\mathfrak{M}$ est l'ensemble des points $(x, y)$ vérifiant l'équation* (1').

*Alors, quel que soit $b_0 > 0$, il existe dans $\Omega$ un rectangle*

$$\Delta = \{|x_j - x_j^0| < a, \ j = 1, \ldots, n, \ |y - y^0| < b\}, \quad b < b_0, \tag{3'}$$

*tel que l'ensemble $\mathfrak{M} \cap \Delta$ soit défini par une fonction continûment dérivable:*

$$y = \psi(x) = \psi(x_1, \ldots, x_n), \quad x \in \Delta^0, \tag{4'}$$

$$\Delta^0 = \{|x_j - x_j^0| < a, \ j = 1, \ldots, n\}. \tag{5'}$$

Les dérivées partielles de la fonction $\psi$ sont données par la formule

$$\frac{\partial \psi}{\partial x_j} = -\frac{\partial f}{\partial x_j} \Big/ \frac{\partial f}{\partial y} \quad (j = 1, \ldots, n). \tag{10'}$$

Si la fonction $f$ (dans le cas des théorèmes 1 et 1') possède des dérivées partielles continues d'ordre supérieur $l$, alors la fonction

implicite possède aussi des dérivées d'ordre $l$ que l'on peut trouver en dérivant $l$ fois l'expression (10) ou (10').

Exemple. Supposons que la fonction $f(x, y)$ envisagée dans le théorème 1 possède des dérivées partielles secondes continues. En dérivant, par exemple l'expression (10), on obtient

$$\psi''(x) = -\frac{f'_y(f''_{x^2}+f''_{xy}\psi')-f'_x(f''_{xy}+f''_{y^2}\psi')}{(f'_y)^2}.$$

## § 16. Plan tangent et normale

Soit une surface $S$ définie par l'équation

$$F(x, y, z) = 0. \tag{1}$$

On admettra que $F(x_0, y_0, z_0) = 0$ et que dans un voisinage du point $(x_0, y_0, z_0)$ la fonction $F$ possède des dérivées partielles continues non toutes nulles. Alors

$$\text{grad}_0 F = ((F'_x)_0, (F'_y)_0, (F'_z)_0) \neq 0. \tag{2}$$

On écrit $(\Phi)_0$ pour $\Phi(x_0, y_0, z_0)$.

Supposons pour fixer les idées que $(F'_z)_0 \neq 0$. Alors le théorème des fonctions implicites nous dit qu'il existe un voisinage du point $(x_0, y_0, z_0)$ dans lequel la surface $S$ est définie par une fonction continûment dérivable $z = f(x, y)$. On sait que le plan tangent à $S$ en $(x_0, y_0, z_0)$ a pour équation

$$z - z_0 = (f'_x)_0 (x - x_0) + (f'_y)_0 (y - y_0),$$

où

$$(f'_x)_0 = -(F'_x)_0/(F'_z)_0, \quad (f'_y)_0 = -(F'_y)_0/(F'_z)_0.$$

L'équation du plan tangent à $S$ en $(x_0, y_0, z_0)$ s'écrit maintenant:

$$(F'_x)_0 (x - x_0) + (F'_y)_0 (y - y_0) + (F'_z)_0 (z - z_0) = 0, \tag{3}$$

celle de la normale à $S$ en $(x_0, y_0, z_0)$ est

$$\frac{x-x_0}{(F'_x)_0} = \frac{y-y_0}{(F'_y)_0} = \frac{z-z_0}{(F'_z)_0}. \tag{4}$$

On obtient les mêmes équations (3) et (4) en admettant que $(F'_x)_0 \neq 0$ ou $(F'_y)_0 \neq 0$. Dans ces cas, au voisinage du point $(x_0, y_0, z_0)$ la surface $S$ est définie par les équations

$$x = \varphi(y, z), \quad y = \psi(x, z).$$

On voit que, si la condition (2) est réalisée, un morceau de surface $S$ appartenant à un voisinage assez petit de $(x_0, y_0, z_0)$ possède en chacun de ses points un plan tangent qui varie continûment lorsque $(x_0, y_0, z_0)$ se déplace continûment. Un tel morceau s'appelle *morceau différentiable de la surface S*.

Autre chose si $\text{grad}_0 F = 0$. La surface $S$ dans ce cas peut admettre ou non un plan tangent en $(x_0, y_0, z_0)$.

EXEMPLE. L'équation

$$z^2 + y^2 - x^2 = 0 \tag{5}$$

représente un cône circulaire d'axe $Ox$ et de sommet $O$ (fig. 102).

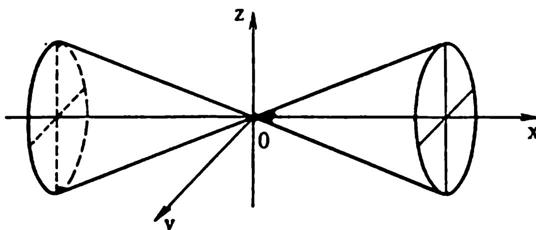



Fig. 102

Le premier membre de l'équation (5) possède des dérivées partielles

$$F'_x = -2x, \ F'_y = 2y, \ F'_z = 2z,$$

non toutes nulles si $(x, y, z) \neq (0, 0, 0)$. Dans un tel point, que nous désignerons par $(x_0, y_0, z_0)$, le plan tangent est défini par l'équation

$$-x_0 (x - x_0) + y_0 (y - y_0) + z_0 (z - z_0) = 0.$$

Le cône envisagé ne possède pas de plan tangent en l'origine des coordonnées. Dans ce cas $\text{grad}_0 F = 0$.

Les points $(x_0, y_0, z_0)$ de $S$ en lesquels $\text{grad}_0 F = 0$ s'appellent *points singuliers* de $S$.

Considérons une fonction continûment dérivable

$$u = f(x, y, z) \tag{6}$$

sur un domaine $\Omega$ de points $(x, y, z)$. Supposons que

$$f(x_0, y_0, z_0) = A, \quad (x_0, y_0, z_0) \in \Omega.$$

Si les dérivées partielles de $f$ en $(x_0, y_0, z_0)$ ne sont pas toutes nulles, alors l'équation $f(x, y, z) = A$ définit au voisinage de ce point une surface différentiable $S$ appelée *surface de niveau de la fonction* (6).

Le plan tangent à $S$ en $(x_0, y_0, z_0)$ a pour équation

$$\left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0) + \left(\frac{\partial f}{\partial z}\right)_0 (z - z_0) = 0.$$

La normale à $S$ en $(x_0, y_0, z_0)$, c'est-à-dire la droite perpendiculaire au plan tangent et passant par $(x_0, y_0, z_0)$, est visiblement

représentée par l'équation

$$\frac{X-x_0}{\left(\frac{\partial f}{\partial x}\right)_0}=\frac{Y-y_0}{\left(\frac{\partial f}{\partial y}\right)_0}=\frac{Z-z_0}{\left(\frac{\partial f}{\partial z}\right)_0}.$$

On remarque que le vecteur

$$\operatorname{grad} f_0=\left(\left(\frac{\partial f}{\partial x}\right)_0,\ \left(\frac{\partial f}{\partial y}\right)_0;\ \left(\frac{\partial f}{\partial z}\right)_0\right)$$

est dirigé suivant la normale à la surface $S$.

L'équation $z = f(x, y)$, où la fonction $f$ possède des dérivées partielles continues, définit une surface différentiable $S$. Posons $f(x_0, y_0) = A$. Si les dérivées partielles $\left(\frac{\partial f}{\partial x}\right)_0$ et $\left(\frac{\partial f}{\partial y}\right)_0$ ne sont pas toutes nulles en $(x_0, y_0)$, alors l'équation $f(x, y) = A$ représente une courbe différentiable au voisinage de ce point, appelée *courbe de niveau* de la fonction $z = f(x, y)$.

L'équation de la tangente à $\Gamma$ en $(x_0, y_0)$ est

$$\left(\frac{\partial f}{\partial x}\right)_0(x-x_0)+\left(\frac{\partial f}{\partial y}\right)_0(y-y_0)=0.$$

Le vecteur $\left(\left(\frac{\partial f}{\partial x}\right)_0,\ \left(\frac{\partial f}{\partial y}\right)_0\right)$ est dirigé suivant la normale à $\Gamma$ en $(x_0, y_0)$.

## § 17. Systèmes de fonctions implicites

Nous avons examiné plus haut le problème de l'existence d'une fonction implicite, continue et dérivable, définie par une équation.

Nous allons envisager maintenant le même problème pour un ensemble de fonctions implicites $y_1, \ldots, y_m$, définies par un système d'équations

$$\left\{\begin{array}{l} f_1(x_1, \ldots, x_n;\ y_1, \ldots, y_m)=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_m(x_1, \ldots, x_n;\ y_1, \ldots, y_m)=0. \end{array}\right. \tag{1}$$

Ainsi, les fonctions $y_1, \ldots, y_m$ cherchées sont solutions du système (1): $y_i = \psi_i(x_1, \ldots, x_n)$ $(i = 1, \ldots, m)$.

Etablissons les conditions d'existence de la solution du système (1) et celles de dérivabilité des fonctions $y_i$.

THÉORÈME. *Soit donné un système d'équations* (1).

*Supposons que les fonctions $f_j$ sont définies dans un voisinage (à $n+m$ dimensions) $\Omega$ d'un point $(x^0, y^0) = (x_1^0, \ldots, x_n^0;\ y_1^0, \ldots, y_m^0)$ de l'espace $R_{n+m}$ des points $(x, y) = (x_1, \ldots, x_n;\ y_1, \ldots, y_m)$, continues dans $\Omega$ avec leurs dérivées partielles premières.*

*Supposons de plus que le déterminant de Jacobi*[1] *(ou jacobien)*

$$\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \dots & \frac{\partial f_1}{\partial y_m} \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial y_1} & \dots & \frac{\partial f_m}{\partial y_m} \end{vmatrix} = \left| \frac{\partial f_i}{\partial y_j} \right| = \frac{D(f_1, \dots, f_m)}{D(y_1, \dots, y_m)} \neq 0 \qquad (2)$$

*en* $(x^0, y^0)$ *et que le point* $(x^0, y^0)$ *vérifie le système* (1).

*Soit* $\mathfrak{M}$ *l'ensemble de tous les points* $(x, y)$ *vérifiant le système* (1).

*Sous ces conditions, quel que soit* $b_0 > 0$, *il existe un rectangle*

$$\Delta = \{|x_j - x_j^0| < a \ (j = 1, \dots, n), \\ |y_i - y_i^0| < b \ (i = 1, \dots, m)\}, \ b < b_0, \qquad (3)$$

*contenu dans* $\Omega$ *et tel que l'ensemble* $\mathfrak{M} \cap \Delta$ *est défini par les fonctions continûment dérivables*

$$y_i = \psi_i(x) \ (i = 1, \dots, m), \ x \in \Delta^0, \qquad (4)$$

$$\Delta^0 = \{|x_j - x_j^0| < a, \ j = 1, \dots, n\}. \qquad (5)$$

En d'autres termes, le rectangle est tel que sur sa projection $\Delta^0$ sur le sous-espace engendré par les coordonnées $(x_1, \dots, x_n)$ on peut définir des fonctions continûment dérivables (4) vérifiant les équations (1):

$$f_j(x, \psi_1(x), \dots, \psi_m(x)) \equiv 0, \quad x \in \Delta^0 \ (j = 1, \dots, m),$$

et les inégalités $|\psi_j(x) - y_j^0| < b$. Ces fonctions sont uniques en ce sens que les coordonnées de tout point $(x, y) \in \mathfrak{M} \cap \Delta$ vérifient les équations (4).

En particulier, $y_j^0 = \psi_j(x^0)$ $(j = 1, \dots, m)$, car $(x^0, y^0) \in \mathfrak{M} \cap \Delta$.

REMARQUE 1. On peut admettre dans le théorème que le rectangle $\Delta$ et sa projection $\Delta^0$ sont définis par

$$\Delta = \{|x_j - x_j^0| < a_j \quad (j = 1, \dots, n); \\ |y_i - y_i^0| < b_i \quad (i = 1, \dots, m)\}, \qquad (3^*)$$

$$\Delta^0 = \{|x_j - x_j^0| < a_j, \ j = 1, \dots, n\}, \qquad (5^*)$$

les nombres $a_j$ et $b_i$ étant en général distincts. En effet, si le théorème est vrai pour le rectangle (3*) pour certains $a_j$ et $b_i$, alors en posant $b = \min b_i$ on peut, en raison de la continuité des fonctions $\psi_i$, exhiber un nombre $a < a_j$ $(j = 1, \dots, n)$ tel que les points $(x, \psi_1(x), \dots, \psi_m(x))$, où $x \in \{|x_j - x_j^0| < a, \quad j = 1, \dots, n\}$, soient contenus dans le rectangle (3).

---

[1]) Carl Jacobi, mathématicien allemand (1804-1851).

A signaler toutefois qu'on ne peut faire en sorte que $a = b$ dans (3), à preuve l'exemple de l'équation $F(x, y) = y - x^2 = 0$, $x_0 = y_0 = 0$.

Prouvons le théorème ci-dessus pour le cas particulier de deux équations

$$\begin{cases} f_1(x_1, x_2, y_1, y_2) = 0, \\ f_2(x_1, x_2, y_1, y_2) = 0. \end{cases} \tag{1'}$$

Il nous faut prouver que si les fonctions $f_1$ et $f_2$ sont continûment dérivables dans un voisinage d'un point $M^0 = (x_1^0, x_2^0, y_1^0, y_2^0) \in R_4$ vérifiant les équations (1') et que le jacobien

$$\begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} \end{vmatrix} \neq 0 \tag{2'}$$

en $M^0$, alors pour tout $b_0 > 0$ il existe un rectangle

$$\Delta = \{|x_1 - x_1^0| < a, |x_2 - x_2^0| < a, |y_1 - y_1^0| < b, |y_2 - y_2^0| < b\} \quad (b < b_0), \tag{3'}$$

contenu dans le voisinage indiqué, et des fonctions continûment dérivables

$$\begin{cases} y_1 = \psi_1(x_1, x_2), \\ y_2 = \psi_2(x_1, x_2), \end{cases} \quad (x_1, x_2) \in \Delta^0, \tag{4'}$$

définies sur la projection

$$\Delta^0 = \{|x_1 - x_1^0| < a, |x_2 - x_2^0| < a\} \tag{5'}$$

de $\Delta$, satisfaisant les équations (1) et telles que

$$y_1^0 = \psi_1(x_1^0, x_2^0), \quad y_2^0 = \psi_2(x_1^0, x_2^0).$$

De plus, pour $(x_1, x_2) \in \Delta^0$,

$$(x_1, x_2, \psi_1(x_1, x_2), \psi_2(x_1, x_2)) \in \Delta. \tag{6}$$

Les fonctions $\psi_1$ et $\psi_2$ sont les seules à donner toutes les solutions des équations (1') dans le rectangle $\Delta$, autrement dit, si un point $(x_1, x_2, y_1, y_2) \in \Delta$ vérifie les équations (1'), alors ses coordonnées sont reliées par les relations (4').

Si le jacobien (2') n'est pas nul en $M^0$, c'est que l'un de ses éléments ne l'est pas. On peut, sans nuire à la généralité, admettre que

$$\frac{\partial f_1}{\partial y_1} \neq 0. \tag{7}$$

On peut toujours se ramener à ce cas en renumérotant au besoin $f_1$, $f_2$ et $y_1$, $y_2$.

Comme les dérivées partielles de $f_1$ et de $f_2$ sont continues par hypothèse, il existe un voisinage assez petit du point $M^0$ dans lequel le jacobien (2') et la dérivée $\frac{\partial f_1}{\partial y_1}$ sont différents de 0. Donc, les conditions du théorème 1' du § 15 sont réunies pour la première équation (1') considérée par rapport à la fonction inconnue $y_1(x_1, x_2, y_2)$. Par suite, pour tout $b_0$ il existe un rectangle

$$\Delta_1 = \{|x_1 - x_1^0| < \alpha, |x_2 - x_2^0| < \alpha, |y_2 - y_2^0| < \beta, |y_1 - y_1^0| < \gamma\} \quad (\gamma < b_0) \tag{8}$$

et une fonction continûment dérivable

$$y_1 = \varphi(x_1, x_2, y_2), \tag{9}$$

$$(x_1, x_2, y_2) \in \Delta_1^0 = \{|x_1 - x_1^0| < \alpha, |x_2 - x_2^0| < \alpha, |y_2 - y_2^0| < \beta\},$$

vérifiant la première équation (1′)

$$f_1(x_1, x_2, \varphi(x_1, x_2, y_2), y_2) = 0, \tag{10}$$

où

$$(x_1, x_2, y_2) \in \Delta_1^0, \ (x_1, x_2, \varphi(x_1, x_2, y_2), y_2) \in \Delta_1. \tag{11}$$

Ceci étant, la fonction $\varphi$ est unique en ce sens que les coordonnées de tout point $(x_1, x_2, y_1, y_2) \in \Delta_1$ vérifiant la première équation (1′) sont reliées par la relation (9); en particulier

$$y_1^0 = \varphi(x_1^0, x_2^0, y_2^0). \tag{12}$$

REMARQUE 2. A noter que dans (8) on aurait pu dans un premier temps admettre que $\alpha = \beta$. Mais dans la suite il aurait fallu réduire les nombres $\alpha$ et $\beta$ de façon non proportionnelle. Les nombres réduits $\alpha$ et $\beta$ conviennent aussi pour la première étape des raisonnements.

Ainsi, nous avons obtenu l'identité (10) qui est vraie quels que soient $(x_1, x_2, y_2) \in \Delta_1^0$. Cette identité reste en vigueur si l'on admet que $y_2$ est une fonction continûment dérivable $y_2 = \psi_2(x_1, x_2)$, telle que

$$(x_1, x_2, \psi_2(x_1, x_2)) \in \Delta_1^0. \tag{13}$$

Or, il existe une infinité de telles fonctions $\psi_2$. Notre but est d'en choisir une telle que les fonctions

$$\begin{cases} y_1 = \varphi(x_1, x_2, \psi_2(x_1, x_2)) = \psi_1(x_1, x_2), \\ y_2 = \psi_2(x_1, x_2) \end{cases} \tag{14}$$

vérifient identiquement la deuxième équation (1′).

Portons la fonction trouvée $\varphi$ dans la deuxième équation (1′):

$$f_2(x_1, x_2, \varphi(x_1, x_2, y_2), y_2) = 0 \tag{15}$$

et cherchons la fonction $y_2(x_1, x_2)$ vérifiant cette équation. Posons

$$\Phi(x_1, x_2, y_2) = f_2(x_1, x_2, \varphi(x_1, x_2, y_2), y_2).$$

La fonction $\Phi$ est continûment dérivable pour tout $(x_1, x_2, y_2) \in \Delta_1^0$ (cf. (11)). Elle satisfait les égalités

$$\Phi(x_1^0, x_2^0, y_2^0) = f_2(x_1^0, x_2^0, \varphi(x_1^0, x_2^0, y_2^0), y_2^0) = f_2(x_1^0, x_2^0, y_1^0, y_2^0) = 0$$

(cf. hypothèse du théorème et (12)). De plus,

$$\frac{\partial \Phi}{\partial y_2} \neq 0.$$

En effet, on a pour les points $(x_1, x_2, y_2) \in \Delta_1^0$:

$$\frac{\partial \Phi}{\partial y_2} = \frac{\partial f_2}{\partial y_1}\frac{\partial \varphi}{\partial y_2} + \frac{\partial f_2}{\partial y_2} = \frac{\partial f_2}{\partial y_1}\left(-\frac{\partial f_1}{\partial y_2}\Big/\frac{\partial f_1}{\partial y_1}\right) + \frac{\partial f_2}{\partial y_2} =$$

$$= \left(\frac{\partial f_1}{\partial y_1}\frac{\partial f_2}{\partial y_2} - \frac{\partial f_2}{\partial y_1}\frac{\partial f_1}{\partial y_2}\right)\Big/\frac{\partial f_1}{\partial y_1} = \begin{vmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} \end{vmatrix} \Big/ \frac{\partial f_1}{\partial y_1} \neq 0. \tag{16}$$

Dans la première égalité (16) on a appliqué la règle de dérivation d'une fonction composée; dans la deuxième on a tenu compte de ce que

$$\frac{\partial\varphi}{\partial y_2} = -\frac{\partial f_1}{\partial y_2}\bigg/\frac{\partial f_1}{\partial y_1}$$

en vertu de (10). Les dérivées partielles de $f_1$ et de $f_2$ sont calculées aux points $x_1 = x_1$, $x_2 = x_2$, $y_1 = \varphi(x_1, x_2, y_2)$, $y_2 = y_2$. Dans la dernière égalité, il faut tenir compte de (2') et de (7).

On constate que le premier membre de l'équation (15) vérifie toutes les conditions du théorème 1' du § 15. Donc, dans le rectangle $\Delta_1^0$ (cf. (9)) il existe un rectangle plus petit que nous désignerons encore par $\Delta_1^0$ (cf. remarque 2 plus haut) et une fonction continûment dérivable

$$y_2 = \psi_2(x_1, x_2),$$
$$(x_1, x_2) \in \Delta^0 = \{|x_1 - x_1^0| < \alpha,\ |x_2 - x_2^0| < \alpha\},$$
$$(x_1, x_2, \varphi_2(x_1, x_2)) \in \Delta_1^0,$$

satisfaisant l'équation (15):

$$\Phi(x_1, x_2, \psi_2(x_1, x_2)) = f_2(x_1, x_2, \varphi(x_1, x_2, \psi_2(x_1, x_2)), \psi_2(x_1, x_2)) =$$
$$= f_2(x_1, x_2, \psi_1(x_1, x_2), \psi_2(x_1, x_2)) = 0,$$

$$(x_1, x_2) \in \Delta^0, \quad (x_1, x_2, \psi_2(x_1, x_2)) \in \Delta_1^0, \quad \psi_2(x_1^0, x_2^0) = y_2^0.$$

Ceci étant, les coordonnées de tout point $(x_1, x_2, y_2) \in \Delta_1^0$ vérifiant l'équation (15) sont reliées par la relation $y_2 = \psi_2(x_1, x_2)$. Mais alors (cf. (11))

$$(x_1, x_2, \psi_1(x_1, x_2), \psi_2(x_1, x_2)) \in \Delta_1.$$

On a ainsi prouvé l'existence de fonctions continûment dérivables

$$\begin{cases} y_1 = \psi_1(x_1, x_2), \\ y_2 = \psi_2(x_1, x_2), \end{cases} \quad (x_1, x_2) \in \Delta^0,$$

satisfaisant les équations (1') et telles que

$$y_1^0 = \psi_1(x_1^0, x_2^0), \quad y_2^0 = \psi_2(x_1^0, x_2^0).$$

Ceci étant,

$$(x_1, x_2, \psi_1(x_1, x_2), \psi_2(x_1, x_2)) \in \Delta_1 \tag{17}$$

et tout point $(x_1, x_2, y_1, y_2) \in \Delta_1$ satisfaisant les équations (1') est de la forme (17).

On passe de $\Delta_1$ à $\Delta$ à l'aide de la remarque 1.

REMARQUE 3. Indiquons une méthode de calcul des dérivées partielles $\frac{\partial y}{\partial x_k}$. Supposons que toutes les conditions du théorème 1 sont remplies. En portant les fonctions $y_j = \psi_j(x)$ dans (1), on obtient le système d'identités:

$$\begin{cases} f_1(x, \psi_1(x), \ldots, \psi_m(x)) \equiv 0, \\ \dots\dots\dots\dots\dots\dots \\ f_m(x_1, \psi_1(x), \ldots, \psi_m(x)) \equiv 0. \end{cases} \tag{18}$$

En dérivant chaque fonction du système (18), on trouve

$$\left\{\begin{array}{l} \frac{\partial f_1}{\partial x_k}+\frac{\partial f_1}{\partial y_1}\frac{\partial \psi_1}{\partial x_k}+\ldots+\frac{\partial f_1}{\partial y_m}\frac{\partial \psi_m}{\partial x_k}\equiv 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{\partial f_m}{\partial x_k}+\frac{\partial f_m}{\partial y_1}\frac{\partial \psi_1}{\partial x_k}+\ldots+\frac{\partial f_m}{\partial y_m}\frac{\partial \psi_m}{\partial x_k}\equiv 0. \end{array}\right. \tag{19}$$

Le système (19) est linéaire en les dérivées

$$\frac{\partial y_j}{\partial x_k}\equiv\frac{\partial \psi_j}{\partial x_k} \qquad (j=1, \ldots, m).$$

Son déterminant qui n'est autre que le jacobien

$$\frac{D(f_1, \ldots, f_m)}{D(y_1, \ldots, y_m)}$$

est différent de 0. Donc, ce système admet une solution unique:

$$\frac{\partial y_1}{\partial x_k}=-\frac{\dfrac{D(f_1, \ldots, f_m)}{D(x_k, y_2, \ldots, y_m)}}{\dfrac{D(f_1, \ldots, f_m)}{D(y_1, \ldots, y_m)}}, \ldots, \frac{\partial y_m}{\partial x_k}=-\frac{\dfrac{D(f_1, \ldots, f_m)}{D(y_1, \ldots, y_{m-1}, x_k)}}{\dfrac{D(f_1, \ldots, f_m)}{D(y_1, \ldots, y_m)}}.$$

## § 18. Applications

Soit donné un système de fonctions continûment dérivables

$$y_j=\varphi_j(x)=\varphi_j(x_1, \ldots, x_n), \quad x\in\Omega \quad (j=1, \ldots, m), \tag{1}$$

où $\Omega$ est un ensemble ouvert de points $x=(x_1, \ldots, x_n)$. On dira que le système (1) définit une *application continûment dérivable*

$$y=Ax, \quad x\in\Omega, \tag{1'}$$

de l'ensemble $\Omega$ sur un ensemble $\Omega'$ de points $y=(y_1, \ldots, y_m)$. On écrira encore $\Omega'=A(\Omega)$ et on dira que $\Omega'$ est l'*image* de $\Omega$ et $\Omega$ l'*image réciproque* de $\Omega'$ par l'application $A$.

Soit $B$ une application continûment dérivable d'un ensemble ouvert $\Lambda$ de points $y$ sur un ensemble de points $z=(z_1, \ldots, z_m)$, définie par le système:

$$z_j=\psi_j(y)=\psi_j(y_1, \ldots, y_m), \quad y\in\Lambda \quad (j=1, \ldots, m).$$

Donc, $z=By$, $y\in\Lambda$.

REMARQUE. Signalons que si $x^0\in\Omega$ et $y^0=Ax^0\in\Lambda$, alors en vertu de la continuité de $A$ il existe un voisinage $V(x^0)$ du point $x^0$ dont l'image par $A$ est contenue dans $\Lambda$. Rétrécissons $\Omega$ jusqu'aux dimensions de $V(x^0)$. On obtient alors que $\Omega'\subset\Lambda$.

Si $\Omega'\subset\Lambda$, on peut alors envisager une application composée continûment dérivable $z=BAx$, $x\in\Omega$, définie par le système $z_j=\psi_j(\varphi_1(x), \ldots, \varphi_m(x))$, $x\in\Omega$ $(j=1, \ldots, m)$.

Les jacobiens des applications $A$, $B$ et $BA$ sont reliés par les égalités remarquables:

$$\frac{D(z_1, \ldots, z_m)}{D(x_1, \ldots, x_m)} = \left|\frac{\partial z_i}{\partial x_j}\right| = \left|\sum_{s=1}^{m} \frac{\partial z_i}{\partial y_s} \frac{\partial y_s}{\partial x_j}\right| =$$

$$= \left|\frac{\partial z_i}{\partial y_s}\right| \cdot \left|\frac{\partial y_s}{\partial x_j}\right| = \frac{D(z_1, \ldots, z_m)}{D(y_1, \ldots, y_m)} \frac{D(y_1, \ldots, y_m)}{D(x_1, \ldots, x_m)} \qquad (2)$$

dont la démonstration, comme on voit, repose sur l'application des règles de dérivation d'une fonction composée et de multiplication des déterminants.

En particulier si $B$ est l'application réciproque de $A$, alors $BA$ est une application identique dont le jacobien est 1, donc

$$1 = \frac{D(x_1, \ldots, x_m)}{D(y_1, \ldots, y_m)} \cdot \frac{D(y_1, \ldots, y_m)}{D(x_1, \ldots, x_m)}, \quad x \in \Omega. \qquad (3)$$

On admettra maintenant que l'application continûment dérivable $y = Ax$ définie par les égalités (1) a un jacobien

$$\frac{D(y_1, \ldots, y_m)}{D(x_1, \ldots, x_m)} \neq 0, \quad x \in \Omega,$$

c'est-à-dire partout non nul sur l'ouvert $\Omega$.

Citons sans les prouver les propriétés suivantes:

1) $\Omega' = A(\Omega)$ *est un ouvert* (avec $\Omega$!).

2) *si* $\Omega$ *est un domaine, il en de même de* $\Omega'$.

3) *l'application* $A$ *est localement bijective, c'est-à-dire que pour tout* $x^0 \in \Omega$ *il existe une boule* $V \subset \Omega$ *centrée en* $x^0$ *telle que la restriction de l'application* $A$ *à* $V$ *soit bijective.*

La propriété 3) affirme seulement la bijectivité locale. La bijectivité globale peut ne pas avoir lieu. Par exemple, la transformation $x = \rho \cos\theta$, $y = \rho \sin\theta$ des coordonnées polaires en cartésiennes, pour $\rho > 0$ et $\theta$ quelconque, est continûment dérivable et possède un jacobien positif égal à $\rho$. C'est une application localement bijective de l'ensemble des points $(\rho, \theta)$ $(\rho > 0, -\infty < \theta < \infty)$ sur celui des points $(x, y)$ non confondus avec l'origine. Mais à tout point $(x, y)$ sont associés un $\rho$ et une infinité de valeurs $\theta$ différant entre elles de $2k\pi$ $(k = \pm 1, \pm 2, \ldots)$.

## § 19. Extrémum lié

Soit la fonction $u = F(x, y) = x^2 + y^2$ dans l'espace $R_2$. Il est immédiat de voir que cette fonction représente le carré de la distance d'un point $P(x, y)$ à l'origine $O$ du système de coordonnées rectangulaires. Elle ne présente pas de maximum dans $R_2$. Mais si on l'envisage uniquement pour les points $(x, y)$ de l'ellipse $G(x, y) \equiv \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ $(b > a)$, il est évident qu'elle passe par son maximum aux points $P_0 = (0, b)$ et $P_1 = (0, -b)$ (fig. 103).

Donc, la fonction $u = F(P)$ considérée sur le plan $R_2$ tout entier ne présente pas de maximum, alors que sur l'ellipse elle en admet deux.

Cette circonstance nous amène à poser le problème de l'extrémum d'une fonction dont les arguments sont assujettis à une contrainte subsidiaire.

Soit donc donnée une fonction

$$u = F(P) = F(x_1, \ldots, x_n, y_1, \ldots, y_m)$$

de $n + m$ variables. On demande de trouver l'extrémum de $F(P)$ sachant que ses variables $x_i$ et $y_k$ sont liées par les relations

$$\begin{cases} G_1(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ G_m(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0, \end{cases} \qquad (1)$$

appelées *contraintes.*

Les équations (1) définissent une *surface* dans l'espace $R_{n+m}$.



Fig. 103

DÉFINITION. *On dira qu'un point $P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$ vérifiant les équations* (1) *est un point de maximum (resp. minimum) lié local si dans $R_{n+m}$ il existe un voisinage de $P^0$ tel que, pour tout point $P$ de ce voisinage vérifiant les contraintes* (1), *l'on ait*

$$F(P) \leqslant F(P^0) \quad (resp.\ F(P) \geqslant F(P^0)).$$

Un point de maximum ou de minimum lié local est appelé *point d'extrémum lié local.*

Dans l'exemple envisagé plus haut, le point $P_0 = (0, b)$ est un point de maximum lié local, car pour tous les points $P$ de l'ellipse $u(P) \leqslant u(P_0)$.

Etudions d'abord les conditions nécessaires pour que le point $P^0$ soit un point d'extrémum lié local. Supposons que $P^0$ est un tel point que les fonctions $G_1, \ldots, G_m$ possèdent des dérivées partielles continues et que le jacobien

$$\frac{D(G_1, \ldots, G_m)}{D(y_1, \ldots, y_m)} \neq 0 \qquad (2)$$

au voisinage de $P^0$.

On sait que sous ces conditions le système (1) admet une solution unique au voisinage de $P^0$:

$$y_i = \varphi_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, m),$$

où les fonctions $\varphi_i(x_1, \ldots, x_n)$ possèdent des dérivées partielles continues au point $M^0 = (x_1^0, \ldots, x_n^0)$.

En portant ces fonctions dans $F$ on constate que $F$ dépend uniquement des variables indépendantes $x_1, \ldots, x_n$:

$$F(x_1, \ldots, x_n, \varphi_1(x_1, \ldots, x_n), \ldots, \varphi_m(x_1, \ldots, x_n)) \equiv \equiv \Phi(x_1, \ldots, x_n). \quad (3)$$

Il est évident que si $F$ passe par un extrémum lié local en $P^0$, alors $\Phi(x_1, \ldots, x_n)$ passe en $M^0 = (x_1^0, \ldots, x_n^0)$ par un extrémum local ordinaire dit encore *extrémum local absolu* et réciproquement.

Or. nous savons que

$$\frac{\partial \Phi(M^0)}{\partial x_i} = 0 \quad (i = 1, \ldots, n) \quad \text{ou} \quad d\Phi(M^0) = \sum_{i=1}^{n} \frac{\partial \Phi}{\partial x_i} dx_i = 0, \quad (4)$$

où $dx_i$ $(i = 1, \ldots, n)$ sont les différentielles des variables indépendantes.

Le point $P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$, pour lequel sont réalisées (4) en vertu de (1) (ou de (3)), s'appelle *point stationnaire de la fonction F sous les contraintes* (1).

Nous avons ainsi prouvé qu'*une condition nécessaire pour qu'un point $P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$ soit un point d'extrémum lié local est qu'il soit un point stationnaire de F sous les contraintes* (1).

On se propose maintenant de trouver un point stationnaire sans passer par la résolution du système (1) par rapport aux variables $y_1, \ldots, y_m$, mais en admettant l'existence des fonctions $\varphi_1, \ldots, \varphi_m$. On écrira $(F)_0$ et $(\varphi_i)_0$ pour $F(P^0)$ et $\varphi_i(M^0)$.

La forme de la différentielle première étant intrinsèque, les conditions (4) équivalent à

$$d\Phi(M^0) = dF(P^0) = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial x_i}\right)_0 dx_i + \sum_{k=1}^{m} \left(\frac{\partial F}{\partial y_k}\right)_0 dy_k = 0, \quad (5)$$

où les différentielles dépendantes $dy_1, \ldots, dy_m$ figurant dans $dF$ sont respectivement égales à

$$dy_k = \sum_{i=1}^{n} \left(\frac{\partial \varphi_k}{\partial x_i}\right)_0 dx_i \quad (k = 1, \ldots, m).$$

Ces différentielles et les différentielles indépendantes $dx_1, \ldots, dx_n$ sont reliées par les relations

$$dG_i(P^0) = \sum_{j=1}^{n} \left(\frac{\partial G_i}{\partial x_j}\right)_0 dx_j + \sum_{k=1}^{m} \left(\frac{\partial G_i}{\partial y_k}\right)_0 dy_k = 0 \quad (i = 1, \ldots, m) \quad (6)$$

qui se déduisent à partir des contraintes.

Ainsi, un point stationnaire de la fonction $F$ sous les contraintes (1) peut être aussi défini comme un point $P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$ vérifiant les équations (1), pour lequel sont réalisées les égalités (5), quels que soient les $dx_1, \ldots, dx_n, dy_1, \ldots, dy_m$ reliés par (6).

Soient les vecteurs à $n + m$ dimensions

$$\operatorname{grad}_0 G_j = \operatorname{grad} G_j(P^0) =$$

$$= \left( \left(\frac{\partial G_j}{\partial x_1}\right)_0, \ldots, \left(\frac{\partial G_j}{\partial x_n}\right)_0, \left(\frac{\partial G_j}{\partial y_1}\right)_0, \ldots, \left(\frac{\partial G_j}{\partial y_m}\right)_0 \right) \quad (j = 1, \ldots, m),$$

$$\operatorname{grad}_0 F = \operatorname{grad} F(P^0) =$$

$$= \left( \left(\frac{\partial F}{\partial x_1}\right)_0, \ldots, \left(\frac{\partial F}{\partial x_n}\right)_0, \left(\frac{\partial F}{\partial y_1}\right)_0, \ldots, \left(\frac{\partial F}{\partial y_m}\right)_0 \right),$$

$$dz = (dx_1, \ldots, dx_n, dy_1, \ldots, dy_m).$$

Les équations (5) et (6) peuvent être représentées par les produits scalaires:

$$(\operatorname{grad}_0 F, dz) = 0, \tag{5'}$$

$$(\operatorname{grad}_0 G_j, dz) = 0 \quad (j = 1, \ldots, m). \tag{6'}$$

On obtient donc qu'un point $P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$ est un point stationnaire sous les contraintes (1) si et seulement s'il vérifie les équations (1) et si l'orthogonalité d'un vecteur $dz$ aux gradients $\operatorname{grad}_0 G_1, \ldots, \operatorname{grad}_0 G_m$ entraîne celle de $dz$ à $\operatorname{grad}_0 F$. Mais dans ce cas (voir justifications plus bas) il existe un système unique de nombres $\lambda_1, \ldots, \lambda_m$, tel que

$$\operatorname{grad}_0 F = \sum_{k=1}^{m} \lambda_k \operatorname{grad}_0 G_k. \tag{7}$$

La réciproque est également vraie. Si l'on sait que $\operatorname{grad}_0 F$ peut être pour certains nombres $\lambda_1, \ldots, \lambda_m$ représenté sous la forme (7), c'est-à-dire sous la forme d'une combinaison linéaire des gradients $\operatorname{grad}_0 G_k$ $(k = 1, \ldots, m)$, alors il s'ensuit immédiatement que, dès qu'un vecteur $dz$ est orthogonal aux gradients $\operatorname{grad}_0 G_k$, il l'est automatiquement à $\operatorname{grad}_0 F$.

La preuve de la réciproque est immédiate: de (7) et (6') il résulte

$$(\operatorname{grad}_0 F, dz) = \left(\sum_{k=1}^{m} \lambda_k \operatorname{grad}_0 G_k, dz\right) =$$

$$= \sum_{k=1}^{m} \lambda_k (\operatorname{grad}_0 G_k, dz) = \sum_{k=1}^{m} \lambda_k \cdot 0 = 0.$$

Pour établir la proposition directe on se référera à un théorème d'algèbre linéaire [1]).

Soit $L$ le sous-espace vectoriel de $R_{n+m}$, engendré par les vecteurs $\mathrm{grad}_0\, G_j$ $(j = 1, \ldots, m)$, c'est-à-dire l'ensemble des combinaisons linéaires de la forme (7) correspondant à tous les systèmes de nombres $\lambda_1, \ldots, \lambda_m$. Considérons le sous-espace $L'$ des vecteurs $dz$ orthogonal à $L$, autrement dit, $L'$ est composé de tous les vecteurs $dz$ orthogonaux à $L$ ou, ce qui est équivalent, aux vecteurs $\mathrm{grad}_0 G_j$ $(j = 1, \ldots, m)$. Si [2]) $L'$ est orthogonal à $L$, alors, réciproquement, $L$ est orthogonal à $L'$, c'est-à-dire que $L$ est composé de tous les vecteurs orthogonaux à $L'$. On a vu qu'en un point stationnaire $P^0$, le gradient de $F$ est orthogonal à tous les vecteurs $dz$ orthogonaux aux gradients de $G_j$, i.e. $\mathrm{grad}_0 F$ est orthogonal à $L'$. Donc, $\mathrm{grad}_0 F \in L$ et par suite est une combinaison linéaire de $\mathrm{grad}_0 G_j$, une combinaison qui est unique, car $\mathrm{grad}_0 G_j$ forment un système linéairement indépendant dans $R_{n+m}$. En effet, la matrice

$$\left\| \begin{matrix} \frac{\partial G_1}{\partial x_1} \cdots \frac{\partial G_1}{\partial x_n} & \frac{\partial G_1}{\partial y_1} \cdots \frac{\partial G_1}{\partial y_m} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial G_m}{\partial x_1} \cdots \frac{\partial G_m}{\partial x_n} & \frac{\partial G_m}{\partial y_1} \cdots \frac{\partial G_m}{\partial y_m} \end{matrix} \right\| \qquad (8)$$

est de rang $m$ au voisinage de $P^0$, car on a admis (2). Donc, ses lignes définissent des vecteurs (gradients) formant un système linéairement indépendant [3]).

Il s'ensuit de ce qui précède qu'*un point stationnaire de la fonction $F$ sous les contraintes* (1) *est un point* $P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots \ldots, y_m^0)$ *vérifiant les équations* (1), *pour lequel* $\mathrm{grad}_0 F$ *est une combinaison linéaire de* $\mathrm{grad}_0 G_j$ $(j = 1, \ldots, m)$:

$$\mathrm{grad}_0\, F = \sum_{j=1}^{m} \lambda_j\, \mathrm{grad}_0\, G_j. \qquad (7)$$

On peut encore dire que *pour qu'un point*

$$P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$$

*soit un point stationnaire de la fonction $F$ sous les contraintes* (1), *il est nécessaire et suffisant qu'il existe des nombres $\lambda_1, \ldots, \lambda_m$ réalisant l'égalité* (7).

Comme le rang de la matrice (8) en $P^0$ est égal à $m$, à chaque point stationnaire correspond un système unique de nombres $\lambda_1, \ldots$

---

[1]) Voir chap. 10, § 20, théorèmes 1, 2 et corollaire 1.
[2]) Voir chap. 10, théorème 1 du § 20.
[3]) Ibid. § 14.

..., $\lambda_m$ pour lesquels a lieu l'égalité (7). Cette égalité équivaut à la suivante:

$$\operatorname{grad}_0 [F - \sum_{j=1}^{m} \lambda_j G_j] = 0. \tag{9}$$

La fonction figurant sous le signe du gradient dans (9)

$$L(P, \lambda) = F(P) - \sum_{j=1}^{m} \lambda_j G_j(P), \quad \lambda = (\lambda_1, \ldots, \lambda_m),$$

s'appelle *fonction de Lagrange* et les nombres $\lambda_j$, *multiplicateurs de Lagrange.*

Détaillons les conditions (9):

$$\left\{\begin{aligned} \frac{\partial L}{\partial x_j} &= \frac{\partial F}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial G_i}{\partial x_j} = 0 \quad (j = 1, \ldots, n), \\ \frac{\partial L}{\partial y_k} &= \frac{\partial F}{\partial y_k} - \sum_{i=1}^{m} \lambda_i \frac{\partial G_i}{\partial y_k} = 0 \quad (k = 1, \ldots, m). \end{aligned}\right. \tag{9'}$$

Le problème de la détermination des points stationnaires de $F$ sous les contraintes (1) s'est donc ramené à la résolution d'un système composé des équations (1) et (9').

Récapitulons.

Pour trouver un point stationnaire

$$P^0 = (x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$$

de la fonction $F$ sous les contraintes (1), il faut former la fonction de Lagrange et le système d'équations (9') et résoudre ce système en même temps que les équations (1). On aura en tout $n + 2m$ équations à $n + 2m$ inconnues $x_1, \ldots, x_n, y_1, \ldots, y_m, \lambda_1, \ldots, \lambda_m$. La résolution de ce système par rapport à $x_i$ et $y_j$ nous donnera le point $(x_1^0, \ldots, x_n^0, y_1^0, \ldots, y_m^0)$ stationnaire. Les points d'extrémum lié local se cherchent parmi les points stationnaires. Pour établir si un point stationnaire est un point d'extrémum lié local, il faut étudier le signe de la différentielle seconde de la fonction de Lagrange. Ce faisant, on tiendra compte du fait que les différentielles $dy_k$ dépendent des différentielles $dx_i$.

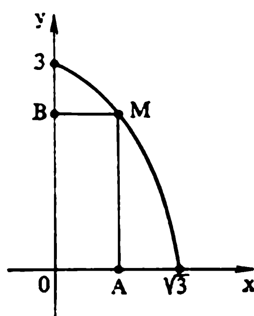


Fig. 104

EXEMPLE. Soit donnée sur le plan $xOy$ une figure limitée par les axes de coordonnées et la parabole $y + x^2 - 3 = 0$ $(0 \leqslant x \leqslant \sqrt{3})$. On demande d'inscrire dans cette figure un rectangle dont les côtés

soient parallèles aux axes de coordonnées, un sommet situé sur la parabole et l'aire maximale (fig. 104).

SOLUTION. Soient $x$ et $y$ les coordonnées du sommet $M$ situé sur la parabole. L'aire du rectangle est $S = xy$. Le point $M$ étant situé sur la parabole, ses coordonnées vérifient l'équation : $y + x^2 - 3 = 0$. Le problème consiste donc à déterminer l'extrémum lié de la fonction $S = xy$ sous la contrainte $y + x^2 - 3 = 0$. Formons la fonction de Lagrange $L(x, y, \lambda) = xy - \lambda(y + x^2 - 3)$. Trouvons les points stationnaires à partir des équations

$$\begin{cases} \dfrac{\partial L}{\partial x} = y - 2\lambda x = 0, \\ \dfrac{\partial L}{\partial y} = x - \lambda = 0, \\ y + x^2 - 3 = 0. \end{cases}$$

La solution de ce système est : $x = 1$, $y = 2$, $\lambda = 1$. Donc le point $(1, 2)$ est un point stationnaire associé au multiplicateur de Lagrange $\lambda = 1$. Etudions la différentielle seconde de la fonction de Lagrange en ce point :

$$L(x, y, 1) = xy - y - x^2 + 3.$$

On a

$$d^2L(x, y, 1) = L''_{x^2}\,dx^2 + 2L''_{xy}dx\,dy + L''_{y^2}dy^2 = 2dx\,(dy - dx).$$

Si l'on admet que $dx$ et $dy$ sont des différentielles de variables indépendantes, alors le signe de $d^2L(x, y, 1)$ n'est pas défini. Or, de $y + x^2 - 3 = 0$ il vient $dy = -2x\,dx$, et au point $(1, 2)$ on a $dy = -2dx$. Donc

$$d^2L(1, 2, 1) = -6dx^2 < 0 \quad (dx \neq 0).$$

Par suite, l'accroissement de la fonction $L(x, y, \lambda)$ au point $x = 1$, $y = 2$, $\lambda = 1$, correspondant à l'accroissement $dx \neq 0$, est $< 0$. Cela veut dire que la fonction $S = xy$ présente en $(1, 2)$ un maximum lié local, car $\Delta S = \Delta L$ sur la parabole $y + x^2 - 3 = 0$.

Donc, le rectangle dont l'aire est maximale est celui dont les côtés sont : $OA = 1$, $OB = 2$.

CHAPITRE 9

# SÉRIES

## § 1. Notion de série

On appelle *série* une expression de la forme

$$u_0 + u_1 + u_2 + \ldots, \tag{1}$$

où les nombres $u_k$, les *termes de la série*, généralement complexes, dépendent des indices $k = 0, 1, 2, \ldots$ On n'associe aucun nombre à cette expression, car la somme d'une infinité de termes n'a pas de sens. La série (1) se note encore:

$$\sum_{k=0}^{\infty} u_k = \sum_{0}^{\infty} u_k. \tag{2}$$

Cette notation est souvent plus commode que (1).

Les nombres

$$S_n = u_0 + u_1 + \ldots + u_n \quad (n = 0, 1, \ldots)$$

s'appellent *sommes partielles de la série* (1).

La série (1) est par définition *convergente* si existe la limite

$$\lim_{n\to\infty} S_n = S.$$

On écrit alors

$$S = u_0 + u_1 + u_2 + \ldots = \sum_{k=0}^{\infty} u_k \tag{3}$$

et on dit que $S$ est la *somme de la série*, c'est-à-dire qu'aux expressions (1) ou (2) on associe le nombre $S$. On dit encore que la série (3) *converge* vers $S$.

REMARQUE. L'égalité $\lim\limits_{n\to\infty} S_n = S$, où $S_n$ et $S$ sont des nombres complexes, se définit comme pour des nombres $S_n$ et $S$ réels, c'est-à-dire qu'elle signifie que pour tout $\varepsilon > 0$ il existe un $N$ tel que $|S_n - S| < \varepsilon$ quel que soit $n > N$. $|S_n - S|$ est le module de la différence des nombres complexes $S_n$ et $S$. Comme pour les

variables réelles, on démontre que la limite de la somme, de la différence, du produit et du quotient de deux variables complexes $u_n$ et $v_n$ est égale respectivement à la somme, à la différence, au produit et au quotient (sous la réserve habituelle que $\lim v_n \neq 0$) des limites de ces variables.

En vertu du critère de Cauchy (qui est valable pour les suites de nombres complexes), *pour que la série* (1) *converge il est nécessaire et suffisant que pour tout* $\varepsilon > 0$ *il existe un* $N$ *tel que pour les entiers naturels* $n > N$ *et pour tout entier* $p$ *l'on ait*

$$|u_{n+1} + \ldots + u_{n+p}| = |S_{n+p} - S_n| < \varepsilon.$$

De là il suit en particulier (si $p = 1$) que si la série (1) converge, son terme général tend vers 0:

$$\lim_{n\to\infty} u_n = 0. \tag{4}$$

La condition (4) est une condition nécessaire mais pas suffisante de convergence, comme nous le verrons dans des exemples ultérieurs.

Considérons encore la série

$$u_{n+1} + u_{n+2} + \ldots = \sum_{k=1}^{\infty} u_{n+k}. \tag{5}$$

Les séries (1) et (5) sont toutes deux convergentes ou divergentes, car le critère de Cauchy se formule de façon analogue pour elles. Si elles convergent, alors la somme de la série (5) est égale à:

$$\lim_{m\to\infty} \sum_{k=1}^{m} u_{n+k} = \lim_{m\to\infty} (S_{n+m} - S_n) = S - S_n.$$

La série (5) s'appelle *reste de la série* (1).

Si les termes de la série (1) sont positifs (donc réels), alors ses sommes partielles forment une suite croissante $S_1 \leqslant S_2 \leqslant S_3 \ldots$ Donc, si cette suite est majorée

$$S_n \leqslant M \quad (n = 1, 2, \ldots),$$

alors la série converge et sa somme vérifie l'inégalité

$$\lim_{n\to\infty} S_n = S \leqslant M.$$

Si elle n'est pas majorée, alors la série diverge:

$$\lim_{n\to\infty} S_n = \infty.$$

On écrit alors

$$\sum_{k=0}^{\infty} u_k = \infty.$$

EXEMPLE. La série

$$1 + z + z^2 + \ldots \quad (z \neq 1) \tag{6}$$

a pour somme partielle $S_n(z) = (1 - z^{n+1})/(1 - z)$. Si $|z| < 1$, alors $|z^{n+1}| = |z|^{n+1}$ et $z^{n+1} \to 0$ $(n \to \infty)$. Donc, la série (6) converge et sa somme est égale à $(1 - z)^{-1}$ sur le disque ouvert $|z| < 1$. Si $|z| \geqslant 1$, alors la série (6) diverge, car son terme général dont le module est supérieur à 1, $|z^n| \geqslant 1$, ne converge pas vers 0 pour $n \to \infty$.

## § 2. Intégrale impropre et série

Considérons l'intégrale

$$\int_a^b f(x)\,dx \tag{1}$$

possédant une seule singularité en $b$. Soit

$$a = b_0 < b_1 < b_2 < \ldots < b, \quad b_k \to b.$$

On peut alors définir la série

$$\int_{b_0}^{b_1} f(x)\,dx + \int_{b_1}^{b_2} f(x)\,dx + \ldots = \sum_{k=0}^{\infty} \int_{b_k}^{b_{k+1}} f\,dx \tag{2}$$

de terme général

$$u_k = \int_{b_k}^{b_{k+1}} f\,dx.$$

THÉORÈME 1. *Si l'intégrale* (1) *converge, il en est de même de la série* (2) *et l'on a l'égalité :*

$$\int_a^b f(x)\,dx = \sum_0^{\infty} \int_{b_k}^{b_{k+1}} f\,dx. \tag{3}$$

En effet,

$$\lim_{n\to\infty} \sum_0^n \int_{b_k}^{b_{k+1}} f\,dx = \lim_{n\to\infty} \int_{b_0}^{b_{n+1}} f\,dx = \int_a^b f\,dx.$$

Si $f$ est positive sur $[a, b[$, alors, réciproquement, la convergence de la série (2) entraîne celle de l'intégrale (1). En effet, supposons que la série converge vers $S$. Pour tout $b'$, où $a < b' < b$, on peut exhiber un $n' = n(b')$ tel que $b' < b_n$, $\forall n > n'$. Donc,

puisque $f(x) \geqslant 0$, on a

$$\int_a^{b'} f\,dx \leqslant \int_a^{b_n} f\,dx = \sum_{k=0}^{n-1} \int_{b_k}^{b_{k+1}} f\,dx \leqslant S,$$

c'est-à-dire que l'intégrale du premier membre est majorée et par suite l'intégrale impropre (1) existe. On a alors l'égalité (3) comme on l'a montré plus haut.

Si la fonction $f$ change de signe sur $[a, b[$, la convergence de la série (2) n'entraîne pas généralement celle de l'intégrale.

Par exemple, la série

$$\sum_{k=0}^{\infty} \int_{2k\pi}^{2(k+1)\pi} \sin x\,dx = \sum_0 0 = 0$$

converge, alors que l'intégrale $\int_0^{\infty} \sin x\,dx$ diverge, car la fonction de $x$

$$\int_0^x \sin t\,dt = 1 - \cos x$$

n'admet pas de limite pour $x \to \infty$.

Théorème 2. *Si une fonction $f(x) \geqslant 0$ est continue et décroissante sur $[0, \infty[$, alors l'intégrale*

$$\int_0^{\infty} f(x)\,dx \tag{3'}$$

*et la série*

$$\sum_{k=0}^{\infty} f(k) = f(0) + f(1) + f(2) + \ldots \tag{4}$$

*sont toutes deux convergentes ou divergentes.*

Démonstration. On a la double inégalité

$$f(k+1) \leqslant \int_k^{k+1} f(x)\,dx \leqslant f(k) \quad (k = 0, 1, \ldots).$$

En sommant par rapport à $k$, on obtient

$$\sum_1^{n+1} f(k) = \sum_0^n f(k+1) \leqslant \int_0^{n+1} f(x)\,dx \leqslant \sum_0^n f(k),$$

d'où la proposition du théorème, vu que les termes de ces relations sont tous monotones croissants pour $n$ croissant.

Du théorème prouvé il s'ensuit que la série

$$1+\frac{1}{2^{\alpha}}+\frac{1}{3^{\alpha}}+\dots \tag{5}$$

converge pour $\alpha>1$ et diverge pour $\alpha\leqslant 1$, car la fonction $\frac{1}{(1+x)^{\alpha}}$, $\alpha>0$, est continue et monotone décroissante sur $[0, \infty[$ et

$$\int_0^{\infty}\frac{dx}{(1+x)^{\alpha}}\begin{cases}<\infty & (\alpha>1),\\ =\infty & (\alpha\leqslant 1).\end{cases}$$

Pour $0<\alpha\leqslant 1$, la série (5) peut servir d'exemple de série divergente dont le terme général ($u_n = n^{-\alpha}$) converge vers 0.

Pour $\alpha\leqslant 0$, on voit immédiatement que la série (5) diverge (son terme général ne converge pas vers 0).

## § 3. Opérations sur les séries

*Si les séries* $\sum\limits_0^{\infty} u_k$ *et* $\sum\limits_0^{\infty} v_k$ *convergent, il en est de même des séries* $\sum\limits_0^{\infty} \alpha u_k$ *($\alpha$ est un nombre),* $\sum\limits_0^{\infty} (u_k\pm v_k)$, *et*

$$\sum_0^{\infty}\alpha u_k=\alpha\sum_0^{\infty}u_k, \tag{1}$$

$$\sum_0^{\infty}(u_k\pm v_k)=\sum_0^{\infty}u_k\pm\sum_0^{\infty}v_k. \tag{2}$$

En effet

$$\sum_0^{\infty}\alpha u_k=\lim_{n\to\infty}\sum_0^{n}\alpha u_k=\alpha\lim_{n\to\infty}\sum_0^{n}u_k=\alpha\sum_0^{\infty}u_k,$$

$$\sum_0^{\infty}(u_k\pm v_k)=\lim_{n\to\infty}\sum_0^{n}(u_k\pm v_k)=$$

$$=\lim_{n\to\infty}\sum_0^{n}u_k\pm\lim_{n\to\infty}\sum_0^{n}v_k=\sum_0^{\infty}u_k\pm\sum_0^{\infty}v_k.$$

Signalons que la convergence de la série du premier membre de (2) n'implique pas nécessairement celle de chacune des séries du

second membre. Par exemple, la série

$$(1-1)+(1-1)+\ldots \tag{3}$$

converge (tous ses termes sont nuls), alors que l'expression $\sum_{0}^{\infty} 1 -$ $- \sum_{0}^{\infty} 1$ n'a pas de sens, car les séries concernées divergent.

Si une série

$$u_0+u_1+u_2+\ldots \tag{4}$$

converge vers $S$, ses termes peuvent être regroupés dans des parenthèses (mais pas permutés), par exemple de la sorte:

$$u_0+(u_1+u_2)+(u_3+u_4+u_5)+\ldots$$

On obtient ainsi une nouvelle série dont les termes sont égaux aux sommes des nombres entre parenthèses. La nouvelle série converge encore vers $S$, car ses sommes partielles forment une suite extraite de la suite convergente des sommes partielles de la série (4).

Il est au contraire généralement illicite d'ouvrir les parenthèses dans une série. Ainsi, en ouvrant les parenthèses dans la série convergente (3), on obtient une série divergente $1-1+1-\ldots$ Cependant, si les parenthèses ne comprennent que des nombres positifs ou négatifs, alors leur ouverture ne modifie ni la convergence ni la somme de la série.

## § 4. Séries à termes positifs

THÉORÈME 1 (CRITÈRE DE COMPARAISON DES SÉRIES). *Soient données deux séries*

$$1)\ \sum_{0}^{\infty} u_k, \qquad 2)\ \sum_{0}^{\infty} v_k$$

*à termes positifs.*

a) *Si* $u_k \leqslant v_k$ $(k=0, 1, \ldots)$, *alors la convergence de la série* 2) *implique celle de la série* 1) *et la divergence de la série* 1) *celle de la série* 2).

b) *Si*

$$\lim_{k\to\infty} \frac{u_k}{v_k} = A > 0, \tag{1}$$

*alors les séries* 1) *et* 2) *sont toutes deux convergentes ou divergentes.*

DÉMONSTRATION. Supposons que la série 2) converge vers $S$. Alors

$$\sum_{0}^{n} u_k \leqslant \sum_{0}^{n} v_k \leqslant S \quad (n=0, 1, \ldots),$$

c'est-à-dire que les sommes partielles de la série 1) sont majorées et la série 1) converge. Sa somme $S'$ est telle que $S' \leqslant S$.

Supposons maintenant que la série 1) diverge: alors (cf. § 1) sa somme partielle croît indéfiniment avec $n$, ce qui en vertu des inégalités

$$\sum_0^n u_h \leqslant \sum_0^n v_k \quad (n=0,\ 1,\ \ldots)$$

entraîne la croissance indéfinie des sommes partielles de la série 2), c'est-à-dire la divergence de 2). Ceci prouve a).

Supposons maintenant que l'on ait (1). Soit un nombre strictement positif $\varepsilon$ tel que $A-\varepsilon>0$. De (1) on déduit la double inégalité

$$A-\varepsilon<\frac{u_k}{v_k}<A+\varepsilon, \quad k>N,$$

valable pour $N$ assez grand, ou la double inégalité

$$(A-\varepsilon)\, v_k < u_k < (A+\varepsilon)\, v_k, \quad k>N. \tag{2}$$

Si la série 2) converge, il en est de même de la série $\sum_{k=N+1}^{\infty} (A+\varepsilon)\, v_k$.

En vertu de la deuxième inégalité (2), la série $\sum_{k=N+1}^{\infty} u_k$ est convergente, donc la série 1). Réciproquement, la convergence de la série 1) entraîne celle de la série $\sum_{k=N+1}^{\infty} (A-\varepsilon)\, v_k$, et par suite celle de la série 2).

On démontre de façon analogue que la divergence de l'une de ces séries implique celle de l'autre. C. q. f. d.

THÉORÈME 2 (CRITÈRE DE D'ALEMBERT [1])). *Soit donnée la série*

$$\sum_0^{\infty} u_k \tag{3}$$

*à termes strictement positifs.*

a) *Si*

$$\frac{u_{k+1}}{u_k} \leqslant q<1 \quad (k=0,\ 1,\ 2,\ \ldots), \tag{4}$$

*alors la série* (3) *converge; si*

$$\frac{u_{k+1}}{u_k} \geqslant 1 \quad (k=0,\ 1,\ 2,\ \ldots), \tag{5}$$

[1]) Jean D'Alembert, mathématicien français (1717-1783).

*elle diverge.*

b) *Si*

$$\lim_{k\to\infty} \frac{u_{k+1}}{u_k} = q, \tag{6}$$

*alors la série* (3) *converge pour* $q < 1$ *et diverge pour* $q > 1$.

DÉMONSTRATION. On a

$$u_n = u_0 \frac{u_1}{u_0} \cdot \frac{u_2}{u_1} \cdots \frac{u_n}{u_{n-1}} \quad (n = 0,\ 1,\ 2,\ \ldots), \tag{7}$$

donc de (4) il résulte que

$$u_n \leqslant u_0 q^n, \quad q < 1 \quad (n = 0,\ 1,\ \ldots).$$

Comme la série $\sum_1^\infty u_0 q^n$ converge, il en est de même de la série (3).
De (5) il s'ensuit que $u_n \geqslant u_0$ ($n = 0,\ 1,\ 2,\ \ldots$), et comme $u_0 > 0$, alors la série (3) diverge (son terme général ne converge pas vers 0).

Si maintenant on a (6) et $q < 1$, alors pour $\varepsilon > 0$ tel que $q + \varepsilon < 1$ on obtient $\frac{u_{k+1}}{u_k} < q + \varepsilon < 1$ ($k \geqslant N$), où $N$ est assez grand. Le critère (4) nous dit alors que la série $\sum_N^\infty u_k$ converge et avec elle la série (3).

Pour $q > 1$, de la propriété (6) il s'ensuit que $\frac{u_{k+1}}{u_k} > 1$ ($k \geqslant N$) pour $N$ assez grand. Le critère (5) nous dit alors que la série $\sum_N^\infty u_k$ diverge et avec elle la série (3).

THÉORÈME 3 (CRITÈRE DE CAUCHY). *Soit donnée une série* (3) *à termes strictement positifs.*

a) *Si*

$$\sqrt[k]{u_k} < q < 1 \quad (k = 0,\ 1,\ \ldots), \tag{8}$$

*alors la série* (3) *converge; si*

$$\sqrt[k]{u_k} \geqslant 1 \quad (k = 0,\ 1,\ \ldots), \tag{9}$$

*alors elle diverge.*

b) *Si*

$$\lim_{k\to\infty} \sqrt[k]{u_k} = q, \tag{10}$$

*alors la série* (3) *converge pour* $q < 1$ *et diverge pour* $q > 1$.

c) *Si*

$$\lim.\ \sup \sqrt[k]{u_k} = q, \tag{11}$$

*alors la série* (3) *converge pour* $q < 1$ *et diverge pour* $q > 1$.

DÉMONSTRATION. De (8) il résulte que $u_k < q^k$ ($q < 1$, $k = 0, 1, \ldots$). Comme la série $\sum_0^\infty q^k$ converge, il en est de même de la série (3). De (9) il s'ensuit que $u_k \geqslant 1$ ($k = 0, 1, \ldots$), c'est-à-dire que la condition nécessaire de convergence n'est pas remplie et la série (3) diverge.

De (10), pour $q < 1$, il s'ensuit que

$$\sqrt[k]{u_k} < q + \varepsilon < 1 \quad (k \geqslant N) \tag{12}$$

pour $N$ assez grand, d'où

$$u_k < (q+\varepsilon)^k \quad (k \geqslant N).$$

Comme la série $\sum_N^\infty (q+\varepsilon)^k$ converge, il en est de même de la série $\sum_N^\infty u_k$ et de la série (3). De (10), pour $q > 1$, il résulte que $\sqrt[k]{u_k} > 1$, c'est-a-dire que $u_k > 1$ ($k \geqslant N$) pour $N$ assez grand, d'où la divergence de la série (3).

De (11) (de même que de (10)), pour $q < 1$, on déduit (12), d'où comme on l'a déjà prouvé, la convergence de la série (3).

Supposons enfin qu'on a (11) pour $q > 1$. Choisissons un nombre $q_1$ tel que $1 < q_1 < q$. En vertu de la propriété de la limite supérieure (cf. chap. 2, § 10) il existe une suite extraite $k_1 < k_2 < \ldots$ telle que

$$\sqrt[k_s]{u_{k_s}} > q_1 > 1 \quad (s = 1, 2, \ldots),$$

i.e.

$$u_{k_s} > q_1^{k_s}.$$

Donc

$$\lim_{s\to\infty} u_{k_s} = \infty$$

et la série (3) diverge.

REMARQUE. La série de terme général $u_n = n^{-\alpha}$ ($\alpha > 0$) converge pour $\alpha > 1$ et diverge pour $\alpha \leqslant 1$ (cf. § 2, (5)).

En outre, dans les deux cas

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = 1, \tag{13}$$

de même que

$$\lim_{n\to\infty} \sqrt[n]{u_n} = 1. \tag{14}$$

Donc, il existe des séries convergentes et des séries divergentes justiciables des critères (13) et (14).

La série $1 + \frac{1}{2} + \frac{1}{3} + \ldots$ s'appelle *série harmonique* (elle diverge, cf. § 2, (5)).

Exemples.

$$1)\ \sum_{0}^{\infty} \frac{x^k}{k!}. \quad 2)\ \sum_{1}^{\infty} \frac{x^k}{k^\alpha}\ (\alpha > 0). \quad 3)\ \sum_{1}^{\infty} (e^{1/k} - 1).$$

$$4)\ \sum_{1}^{\infty} \ln\left(1 + \frac{1}{k}\right). \quad 5)\ \sum_{1}^{\infty} q^{k+\sqrt{k}}\ (q > 0).$$

La série 1) converge pour tout $x \geqslant 0$. Pour $x = 0$, cela est évident ; pour $x > 0$, cela résulte du fait que $\frac{u_{k+1}}{u_k} = \frac{x}{k+1} \to 0$, $k \to \infty$. On sait de plus que cette série est la série de Taylor de la fonction $e^x$ et qu'elle converge vers $e^x$.

La série 2) converge pour $0 \leqslant x < 1$ et diverge pour $x > 1$, car pour $x > 0$ on a $\frac{u_{k+1}}{u_k} = x\left(\frac{k}{k+1}\right)^\alpha \to x$, $k \to \infty$ ; pour $x = 1$, voir remarque ci-dessus. Le cas $x = 0$ est trivial.

Les séries 3) et 4) divergent en vertu du théorème 1 du § 4, car $e^{1/k} - 1 \approx 1/k$ $(k \to \infty)$ et $\ln(1 + (1/k)) \approx 1/k$ $(k \to \infty)$ ( $\approx$ est le symbole de l'égalité asymptotique, cf. chap. 3, §§ 9, 10), et la série harmonique $\sum_{1}^{\infty} k^{-1}$ diverge.

La série 5) converge pour $0 \leqslant q < 1$ et diverge pour $q > 1$, car $\sqrt[k]{u_k} = q^{1+(1/\sqrt{k})} \to q$ $(k \to \infty)$. Pour $q = 1$, elle diverge aussi : son terme général étant égal à 1.

Théorème 4. *Supposons qu'une série*

$$u_0 + u_1 + u_2 + \ldots \tag{15}$$

*à termes positifs converge vers $S$. La série*

$$u'_0 + u'_1 + u'_2 + \ldots \tag{16}$$

*obtenue par une permutation des termes de* (15) *converge aussi vers $S$.*

Démonstration. Soit

$$S'_n = u'_0 + u'_1 + \ldots + u'_n$$

une somme partielle de (16). Les termes de cette somme figurent dans (15) sous d'autres indices, disons $k_0, \ldots, k_n$. Soient $N$ le plus grand d'entre eux et $S_N$ une somme partielle de la série (15). Il est évident que $S'_n \leqslant S_N \leqslant S$. Comme $n$ est arbitraire, la série (16)

converge vers $S' \leqslant S$. Si maintenant on échange les séries (15) et (16) et qu'on refasse le même raisonnement, on obtient $S \leqslant S'$. Donc $S = S'$.

## § 5. Série de Leibniz

On appelle *série de Leibniz* une série de la forme

$$a_0 - a_1 + a_2 - a_3 + \ldots, \tag{1}$$

où les nombres $a_k > 0$, $a_k \geqslant a_{k+1}$; $a_k \to 0$, $k \to \infty$.

Montrons que la série de Leibniz converge vers $S \leqslant a_0$.

En effet, la somme partielle $S_{2n+1}$ d'indice impair $2n + 1$ peut s'écrire sous la forme:

$$S_{2n+1} = a_0 - (a_1 - a_2) - (a_3 - a_4) - \ldots \\ \ldots - (a_{2n-1} - a_{2n}) - a_{2n+1},$$

d'où il résulte de toute évidence qu'elle est majorée par $a_0$:

$$S_{2n+1} \leqslant a_0.$$

Cette somme peut encore s'écrire

$$S_{2n+1} = (a_0 - a_1) + (a_2 - a_3) + \ldots + (a_{2n} - a_{2n+1}),$$

d'où il s'ensuit que $\{S_{2n+1}\}$ est monotone croissante. Dans ce cas, $\lim_{n\to\infty} S_{2n+1} = S \leqslant a_0$. Il est évident aussi que

$$\lim_{n\to\infty} S_{2n} = \lim_{n\to\infty} (S_{2n+1} - a_{2n+1}) = S - 0 = S.$$

Ce qui prouve le théorème.

Exemple. La série $1 - \frac{1}{2} + \frac{1}{3} - \ldots$ est une série de Leibniz. Donc, elle converge et sa somme $S \leqslant 1$ (en fait, $S = \ln 2$, cf. chap. 4, § 16, nº 4).

## § 6. Séries absolument convergentes

On dit qu'une série à termes complexes

$$u_0 + u_1 + u_2 + \ldots \tag{1}$$

est *absolument convergente* si la série

$$|u_0| + |u_1| + |u_2| + \ldots \tag{2}$$

des modules de ses termes converge.

*Une série absolument convergente est convergente.*

En effet, supposons que la série (1) converge absolument; alors la série (2) converge et en vertu du critère de Cauchy pour tout $\varepsilon > 0$ il existe un $N$ tel que $\varepsilon > |u_{n+1}| + \ldots + |u_{n+p}|$ pour tous les $p$ et $n > N$. Alors à fortiori $\varepsilon > |u_{n+1} + \ldots + u_{n+p}|$. Donc, la série (1) converge en vertu du critère de Cauchy.

Les séries convergentes à termes positifs convergent absolument de façon triviale. La série $1 - \frac{1}{2^\alpha} + \frac{1}{3^\alpha} - \ldots$ $(\alpha > 0)$ converge comme série de Leibniz. Mais elle ne converge absolument que pour $\alpha > 1$.

THÉORÈME. *Si une série converge absolument, alors la série obtenue par permutation des termes converge aussi absolument vers la même somme.*

DÉMONSTRATION. Prouvons d'abord le théorème pour le cas où les termes $u_k$ de la série sont des nombres réels.

Posons

$$u_k^+ = \begin{cases} u_k \text{ si } u_k \geqslant 0, \\ 0 \text{ si } u_k < 0, \end{cases} \quad u_k^- = \begin{cases} -u_k \text{ si } u_k \leqslant 0, \\ 0 \text{ si } u_k > 0; \end{cases} \tag{3}$$

les nombres $u_k^+$ et $u_k^-$ sont positifs et

$$u_k = u_k^+ - u_k^-. \tag{4}$$

En plus de la série (1), on étudiera les séries

$$\sum_0^\infty u_k^+ \text{ et } \sum_0^\infty u_k^-. \tag{5}$$

Supposons que la série (1) converge absolument et que ses termes sont des nombres réels $u_k$. Alors les séries (5) convergent aussi, car de toute évidence $u_k^+ \leqslant |u_k|$, $u_k^- \leqslant |u_k|$.

Supposons que la série permutée est de la forme $v_1 + v_2 + v_3 + \ldots$ Introduisons comme plus haut les nombres $v_k^+$ et $v_k^-$. On a (voir justifications plus bas)

$$\sum_0^\infty u_k = \sum_0^\infty (u_k^+ - u_k^-) = \sum_0^\infty u_k^+ - \sum_0^\infty u_k^- =$$
$$= \sum_0^\infty v_k^+ - \sum_0^\infty v_k^- = \sum_0^\infty (v_k^+ - v_k^-) = \sum_0^\infty v_k.$$

La première égalité résulte de (4), la seconde, du § 3, (2) si l'on tient compte de la convergence des séries (5), la troisième, du fait que les séries convergentes à termes positifs sont permutables, la

quatrième, du § 3, (2), la cinquième enfin, du fait que $v_k = v_k^+ - v_k^-$. Le théorème est prouvé pour les $u_k$ réels.

Supposons maintenant que $u_k = \alpha_k + i\beta_k$, les nombres $v_k$ ayant le même sens que précédemment. Comme $|\alpha_k| \leqslant |u_k|$ et $|\beta_k| \leqslant |u_k|$, les séries $\sum_0^\infty \alpha_k$ et $\sum_0^\infty \beta_k$ convergent absolument, et d'après ce qui vient d'être prouvé, leurs termes sont permutables. Donc, si l'on admet que $v_k = \gamma_k + i\delta_k$, on obtient

$$\sum_0^\infty u_k = \sum_0^\infty (\alpha_k + i\beta_k) = \sum_0^\infty \alpha_k + i\sum_0^\infty \beta_k =$$
$$= \sum_0^\infty \gamma_k + i\sum_0^\infty \delta_k = \sum_0^\infty (\gamma_k + i\delta_k) = \sum_0^\infty v_k.$$

C. q. f. d.

## § 7. Séries semi-convergentes et commutativement convergentes à termes réels

On a vu au paragraphe précédent qu'une série absolument convergente à termes réels ou complexes restait absolument convergente et avait la même somme si on changeait l'ordre de ses termes.

Cette propriété de ne pas changer de somme est caractéristique uniquement des séries ab·olument convergentes.

Considérons la série

$$u_0 + u_1 + u_2 + \ldots \tag{1}$$

à termes réels convergente mais pas absolument.

On démontre que quel que soit $S \in [-\infty, \infty]$, il existe une série obtenue par permutation des termes de la série (1) dont la somme est $S$.

On dit qu'une série est *commutativement convergente* si elle converge vers un même nombre quel que soit l'ordre de ses termes. Si une série converge et s'il existe une permutation de ses termes pour laquelle elle n'est plus convergente ou converge vers une autre somme, alors une telle série est dite *semi-convergente*.

De ce qui précède il suit que *pour qu'une série soit absolument convergente, il est nécessaire et suffisant qu'elle soit commutativement convergente.*

Faisons la remarque suivante. Soit donnée une série (1) à termes réels convergente mais pas absolument. La série (1) comporte une infinité de termes positifs et de termes négatifs qui forment séparément des séries divergentes (sinon la série (1) serait absolument convergente).

## § 8. Suites et séries de fonctions. Convergence uniforme

Considérons une suite de fonctions $\{f_k(x)\}$ définies sur un ensemble de points $x = (x_1, \ldots, x_n)$ d'un espace à $n$ dimensions. Ces fonctions sont susceptibles de prendre des valeurs complexes ($f_k(x) = \alpha_k(x) + i\beta_k(x)$). On peut admettre que $x$ est un point complexe ($x = \xi + i\eta$) parcourant un ensemble $E$ de points du plan complexe et $f_k(x)$ des fonctions de la variable complexe $x$.

Supposons que pour toute valeur $x \in E$ la suite $\{f_n(x)\}$ converge vers un nombre $f(x)$ (qui est fonction de $x$).

Par définition, *la suite* $\{f_n(x)\}$ *converge vers* $f(x)$ *uniformément sur* $E$ *s'il existe une suite convergeant vers* 0 *de nombres positifs* $\rho_n$ (*ne dépendant pas de* $x$) *telle que*

$$|f(x) - f_n(x)| \leqslant \rho_n, \quad \forall x \in E. \tag{1}$$

Cette définition équivaut à la suivante: *pour tout* $\varepsilon > 0$ *il existe un* $N$ *tel que pour* $n > N$ *l'on ait*

$$|f(x) - f_n(x)| < \varepsilon, \quad \forall x \in E.$$

En effet, si l'on se place dans le cas de la première définition, pour tout $\varepsilon > 0$ il existe un $N$ tel que

$$\varepsilon \underset{n>N}{>} \rho_n \geqslant |f(x) - f_n(x)|, \quad \forall x \in E,$$

i. e.

$$\varepsilon > |f(x) - f_n(x)|, \quad \forall x \in E, \quad n > N. \tag{2}$$

Inversement, d'après la deuxième définition, pour tout $\varepsilon > 0$ il existe un $N$ tel que l'on ait (2). Donc

$$\varepsilon \geqslant \sup_{x \in E} |f(x) - f_n(x)| = \rho_n, \quad n > N. \tag{3}$$

On voit que les nombres positifs $\rho_n$ ne dépendent pas de $x$ et $|f(x) - f_n(x)| \leqslant \rho_n$, $\rho_n \to 0$, i.e. on retrouve la première définition.

Dans la première définition, pour $\rho_n$ on peut prendre la borne supérieure

$$\sup_{x \in E} |f(x) - f_n(x)| = \rho_n.$$

Si cette borne converge vers 0 pour $n \to \infty$ ($\rho_n \to 0$), alors $\{f_n(x)\}$ converge vers $f(x)$ uniformément sur $E$. La convergence de $\{f_n(x)\}$ vers $f(x)$ n'est pas uniforme si $\rho_n$ ne converge pas vers 0.

On peut donner une troisième définition de la convergence uniforme (au sens de Cauchy): *une suite de fonctions* $\{f_n(x)\}$ *converge*

*uniformément sur $E$ si pour tout $\varepsilon > 0$ on peut exhiber un $N$ tel que*

$$| f_{n+p}(x) - f_n(x) | < \varepsilon \tag{4}$$

*quels que soient $n > N$, $p > 0$ et $x \in E$.*

Si la suite converge uniformément au sens de la deuxième définition, alors pour tout $\varepsilon > 0$ on peut exhiber un $N$ tel que

$$| f_{n+p}(x) - f_n(x) | \leqslant | f_{n+p}(x) - f(x) | + | f(x) - f_n(x) | \leqslant 2\varepsilon$$

quels que soient $n > N$, $p > 0$ et $x \in E$, autrement dit on retrouve la troisième définition. Plaçons-nous maintenant dans le cas de la troisième définition; alors pour chaque valeur $x \in E$ a lieu le critère de Cauchy, donc la suite $\{f_n(x)\}$ converge vers une fonction $f(x)$. Donnons-nous $\varepsilon > 0$ et choisissons $N$ tel que l'indique la troisième définition. Passons à la limite pour $p \to \infty$ dans l'inégalité (4), où $n > N$ est fixé; on obtient $|f(x) - f_n(x)| \leqslant \varepsilon$ $(x \in E)$, d'où

$$\rho_n = \sup_{x \in E} | f(x) - f_n(x) | \leqslant \varepsilon.$$

Comme $n > N$ est arbitraire, il vient $\rho_n \to 0$ $(n \to \infty)$, c'est-à-dire qu'on retrouve la première définition.

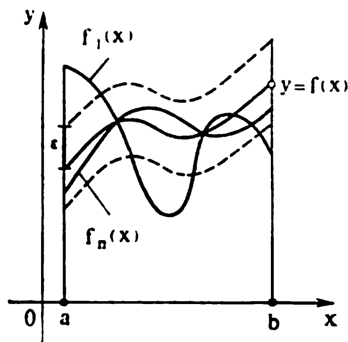



Fig. 105

Traçons le graphe de la fonction $f(x)$ (que nous supposons être continue sur l'intervalle $[a, b]$ (fig. 105). Donnons-nous $\varepsilon > 0$ et comprenons le graphe dans une bande de largeur $2\varepsilon$. Les coordonnées $(x, y)$ de tout point de cette bande vérifient la double inégalité

$$f(x) - \varepsilon < y < f(x) + \varepsilon.$$

Si la suite de fonctions $\{f_n(x)\}$ converge vers $f(x)$ uniformément sur $[a, b]$, alors pour $\varepsilon > 0$ donné, on peut exhiber un $N$ tel que pour tout $n > N$ le graphe de la fonction $y = f_n(x)$ soit compris dans la bande. Si $\{f_n(x)\}$ converge vers $f(x)$ non uniformément sur $[a, b]$, alors on ne peut pas pour tout $\varepsilon > 0$ exhiber un $N$ tel que pour tout $n > N$ tous les graphes de $y = f_n(x)$ soient compris dans une bande de largeur $2\varepsilon$, bien que $f_n(x)$ tende vers $f(x)$ quel que soit $x \in [a, b]$ (cf. exemple 3 plus bas).

Il est immédiat de voir que si $\{f_k(x)\}$ et $\{\varphi_k(x)\}$ sont des suites de fonctions uniformément convergentes sur $E$, alors il en est de même des suites $\{\alpha f_k(x)\}$ ($\alpha$ est une constante) et $\{f_k(x) \pm \varphi_k(x)\}$. Il est de même aisé de voir que si une suite de fonctions converge

uniformément sur $E$, elle convergera uniformément sur $E' \subset E$. La réciproque n'est pas toujours vraie.

Signalons qu'à toute suite de fonctions $\{f_k(x)\}$ est associée la série

$$f_0(x) + [f_1(x) - f_0(x)] + [f_2(x) - f_1(x)] + \ldots$$

dont les sommes partielles à l'ordre $n$ sont respectivement égales à $f_n(x)$.

Soit donnée maintenant une série

$$u_0(x) + u_1(x) + u_2(x) + \ldots \tag{5}$$

dont les termes sont des fonctions complexes de $x \in E$, où $E$ est toujours un ensemble d'un espace à $n$ dimensions ou le plan complexe.

Par définition, *la série* (5) *converge uniformément vers une fonction* $S(x)$ *sur l'ensemble* $E$ si la suite $\{S_k(x)\}$ de ses sommes partielles converge uniformément sur $E$ vers $S(x)$.

On peut notamment énoncer la définition de la convergence uniforme sous la forme suivante: *la série* (5) *converge uniformément sur* $E$ *si pour tout* $\varepsilon > 0$ *on peut exhiber un* $N$ *tel que quels que soient* $n > N$, $p > 0$ *et* $x \in E$ *l'on ait*

$$| u_{n+1}(x) + \ldots + u_{n+p}(x) | < \varepsilon.$$

Le théorème suivant nous fournit un important critère de convergence uniforme d'une série.

THÉORÈME 1 (WEIERSTRASS). *Si les termes d'une série* (5) *vérifient les inégalités*

$$| u_k(x) | \leqslant \alpha_k \quad (k = 0, 1, \ldots), \tag{6}$$

*où* $x \in E$ *et* $\alpha_k$ *sont des nombres ne dépendant pas de* $x$, *et si la série* $\sum_0^\infty \alpha_k$ *converge, alors la série* (5) *converge absolument et uniformément sur l'ensemble* $E$.

En effet, de la convergence de la série $\sum_0^\infty \alpha_k$ et de (6) il s'ensuit que pour tout $\varepsilon > 0$ on peut exhiber un $N$ tel que quels que soient $n > N$, $p > 0$ et $x \in E$, l'on ait

$$\varepsilon > \alpha_{n+1} + \ldots + \alpha_{n+p} \geqslant | u_{n+1}(x) | + \ldots + | u_{n+p}(x) | \geqslant$$
$$\geqslant | u_{n+1}(x) + \ldots + u_{n+p}(x) |.$$

Or, cela signifie que la série (5) converge uniformément sur $E$. La convergence absolue est évidente.

THÉORÈME 2. *Si une suite de fonctions $\{f_n(x)\}$ converge uniformément sur un ensemble E vers une fonction f et si les fonctions $f_n$ sont continues en un point $x^0$, alors il en est de même de f.*

Dans le langage des séries, ce théorème dit que *la somme d'une série uniformément convergente sur E de fonctions continues en un point $x^0 \in E$ est une fonction continue en ce point* [1]).

DÉMONSTRATION. Donnons-nous $\varepsilon > 0$ et choisissons un entier naturel $N$ tel que $|f(x) - f_N(x)| < \varepsilon/3$ pour tout $x \in E$. Ceci est possible puisque $\{f_n\}$ converge uniformément vers $f$. On a d'autre part

$$|f(x) - f(x^0)| \leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(x^0)| + \\ + |f_N(x^0) - f(x^0)| < 2\frac{\varepsilon}{3} + |f_N(x) - f_N(x^0)| \quad (7)$$

pour tout $x \in E$. La fonction $f_N$ étant continue en $x^0$, on peut exhiber un $\delta > 0$ tel que $|f_N(x) - f_N(x^0)| < \varepsilon/3$ pour tous les $x \in E$ pour lesquels $|x - x^0| < \delta$; donc de (7) il s'ensuit que pour de tels $x$ on a

$$|f(x) - f(x^0)| < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

c.q.f.d.

EXEMPLE 1. La série

$$1 + (x - 1) + (x^2 - x) + (x^3 - x^2) + \ldots \quad (8)$$

converge sur l'intervalle [0, 1], mais pas uniformément. Sur l'intervalle $[0, q]$, où $0 < q < 1$, elle est uniformément convergente.

En effet, la somme partielle

$$S_n(x) = x^n \xrightarrow[n \to \infty]{} \begin{cases} 0, & 0 \leqslant x < 1, \\ 1, & x = 1. \end{cases}$$

La valeur absolue du reste $S(x) - S_n(x)$ est égale à

$$|S(x) - S_n(x)| = \begin{cases} x^n, & 0 \leqslant x < 1, \\ 0, & x = 1. \end{cases} \quad (9)$$

Sur l'intervalle $[0, q]$, où $0 < q < 1$,

$$|S(x) - S_n(x)| = x^n \leqslant q^n.$$

Le second membre de cette inégalité ne dépend pas de $x \in [0, q]$ et tend vers 0 pour $n \to \infty$ ($q^n \to 0$). Ceci exprime que la série (8) converge uniformément sur l'intervalle $[0, q]$, où $0 < q < 1$.

D'autre part, on voit sur l'égalité (9) que

$$\sup_{x \in [0, 1]} |S(x) - S_n(x)| = 1.$$

---

[1]) Voir remarque du § 12.

Or, le nombre 1 ne tend pas vers 0 pour $n \to \infty$, donc la série (8) converge sur [0, 1], mais pas uniformément.

EXEMPLE 2. Le terme général de la série

$$\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{3^2} + \ldots \tag{10}$$

vérifie l'inégalité

$$\frac{|\sin nx|}{n^2} \leqslant \frac{1}{n^2}, \quad \forall x \in ]-\infty, \infty[.$$

Or, la série

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

converge. Donc, d'après le théorème de Weierstrass, la série (10) converge uniformément sur $]-\infty, \infty[$.

Comme les termes de la série (10) sont des fonctions continues, d'après le théorème 2 il en est de même de la somme.

EXEMPLE 3. La figure 106 représente le graphe de la fonction $f_n(x)$ $(n = 1, 2, \ldots)$. Cette fonction est linéaire sur chacun des intervalles $[0, 1/n]$, $[1/n, 2/n]$, $[2/n, 1]$. De plus $f_n(0) = f(2/n) = 0$ et $f_n(x) = 0$ sur $[2/n, 1]$, $f_n(1/n) = 1$. Il est évident que

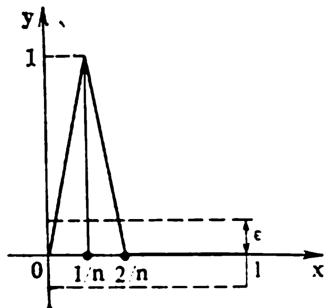


Fig. 106

$$\lim_{n\to\infty} f_n(x) = f(x) = 0, \quad \forall x \in [0, 1],$$

car $f_n(0) = 0 \to 0$ et, si $0 < x \leqslant 1$, alors $f_n(x) = 0$, $\forall n > 2/x$.

Par ailleurs, il est évident que

$$\sup_{x\in[0, 1]} |f_n(x) - f(x)| = \sup_{x\in[0, 1]} f_n(x) = 1.$$

Comme le nombre 1 ne converge pas vers 0 pour $n \to \infty$, la suite $\{f_n(x)\}$ converge vers $f(x)$ sur [0, 1] mais pas uniformément.

Sur la figure 106 est représentée en pointillé une $\varepsilon$-bande de largeur $2\varepsilon$ encadrant la courbe limite $f(x) = 0$ $(0 \leqslant x \leqslant 1)$. Le graphe de la fonction $f_n(x)$ n'est entièrement compris dans cette bande pour aucune valeur de $n$. Ceci n'empêche pas que $f_n(x) \to f(x) = 0$, $\forall x \in [0, 1]$.

Exhibons des critères plus subtils de convergence uniforme des séries, reposant sur l'application de la *transformation d'Abel* (l'analogue de l'intégration par parties) à ces séries.

Soit la série

$$\alpha_0\beta_0 + \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots \tag{11}$$

où $\alpha_k$ et $\beta_k$ sont des fonctions de $x \in E$ (ou des constantes). Posons $B_k = \beta_{n+1} + \beta_{n+2} + \ldots + \beta_{n+k}$ et appliquons la transformation d'Abel à la somme tronquée de la série (11):

$$\sum_{k=1}^{p} \alpha_{n+k}\beta_{n+k} = \alpha_{n+1}\beta_{n+1} + \ldots + \alpha_{n+p}\beta_{n+p} =$$
$$= \alpha_{n+1}B_1 + \alpha_{n+2}(B_2 - B_1) + \ldots + \alpha_{n+p}(B_p - B_{p-1}) =$$
$$= (\alpha_{n+1} - \alpha_{n+2})B_1 + (\alpha_{n+2} - \alpha_{n+3})B_2 + \ldots + (\alpha_{n+p-1} - \alpha_{n+p})B_{p-1} + \alpha_{n+p}B_p =$$
$$= \sum_{k=1}^{p-1} (\alpha_{n+k} - \alpha_{n+k+1})B_k + \alpha_{n+p}B_p. \tag{12}$$

Il est aisé d'établir maintenant deux critères de convergence uniforme (ou simplement de convergence si $\alpha_k$ et $\beta_k$ sont constantes) de la série (11).

THÉORÈME 3 (CRITÈRE DE DIRICHLET DE CONVERGENCE UNIFORME D'UNE SÉRIE). *Si les sommes partielles de la série*

$$\beta_0 + \beta_1 + \beta_2 + \ldots \tag{13}$$

*sont majorées dans leur ensemble et si les fonctions réelles $\alpha_k(x)$ convergent uniformément vers 0 sur $E$ en décroissant, alors la série* (11) *converge uniformément.*

En effet, supposons que les sommes partielles $\sigma_n$ de la série (13) sont inférieures à une constante $M$. Alors pour tous $n$ et $k$

$$|B_k| = |\sigma_{n+k} - \sigma_n| \leqslant |\sigma_{n+k}| + |\sigma_n| \leqslant 2M.$$

Donc, en vertu de (12) et du fait que $\alpha_s$ convergent uniformément vers 0 en décroissant, on a

$$\left|\sum_{1}^{p} \alpha_{n+k}\beta_{n+k}\right| \leqslant 2M \sum_{k=1}^{p-1} (\alpha_{n+k} - \alpha_{n+k+1}) + \alpha_{n+p}2M = 2M\alpha_{n+1} < \varepsilon,$$

quels que soient $n > N$, $p > 0$ et $x \in E$, dès que $N$ est assez grand. Donc, la série (11) converge uniformément. La dernière inégalité est valable pour tous les $x \in E$, car $\alpha_{n+1}(x)$ converge uniformément vers 0.

THÉORÈME 4 (CRITÈRE D'ABEL DE CONVERGENCE UNIFORME D'UNE SÉRIE). *Si des fonctions réelles $\alpha_k$ sont monotones décroissantes et majorées dans leur ensemble et si la série* (13) *converge uniformément sur $E$, alors il en est de même de la série* (11).

En effet, supposons que $|\alpha_k| \leqslant M$ $(k = 0, 1, \ldots)$ (la valeur absolue indique que les fonctions $\alpha_k$ peuvent être négatives). La série (13) étant uniformément convergente, pour tout $\varepsilon > 0$ on peut exhiber un $N$ tel que $|B_k| < \varepsilon$ quels que soient $n > N$ et $k$. Donc, en vertu de (12), et puisque $\alpha_s$ sont

monotones décroissantes, pour tous $n > N$ et $p$, on a

$$\left|\sum_{1}^{p} \alpha_{n+k}\beta_{n+k}\right| \leqslant \varepsilon \sum_{k=1}^{p-1} (\alpha_{n+k}-\alpha_{n+k+1})+\varepsilon\,|\alpha_{n+p}| =$$

$$=\varepsilon\,(\alpha_{n+1}-\alpha_{n+p})+\varepsilon\,|\alpha_{n+p}| \leqslant 3\varepsilon M,$$

i.e. la série (11) est uniformément convergente.

EXEMPLE 4. Les séries

$$\sum_{1}^{\infty} \frac{\cos kx}{k^{\alpha}}, \quad \sum_{1}^{\infty} \frac{\sin kx}{k^{\alpha}} \quad (\alpha > 0) \tag{14}$$

convergent uniformément et absolument sur l'axe réel tout entier pour $\alpha > 1$, car les valeurs absolues de leurs termes de rang $k$ sont inférieures à $k^{-\alpha}$ et la série $\sum k^{-\alpha}$ converge. Le critère de Weierstrass dont on vient de se servir ne passe pas pour $\alpha \leqslant 1$, car la série $\sum k^{-\alpha}$ est alors divergente. Mais pour $0 < \alpha \leqslant 1$, les séries (14) convergent uniformément sur l'intervalle $[\varepsilon, 2\pi - \varepsilon]$ quel que soit $\varepsilon > 0$, où $0 < \varepsilon < \pi$. En effet, les sommes partielles des séries

$$\frac{1}{2} + \cos x + \cos 2x + \cos 3x + \ldots \quad \text{et} \quad \sin x + \sin 2x + \ldots$$

sont égales respectivement à

$$D_n(x) = \frac{\sin\left(n+\frac{1}{2}\right)x}{2\sin\frac{x}{2}}, \quad K_n(x) = \frac{\cos\frac{x}{2} - \cos\left(n+\frac{1}{2}\right)x}{2\sin\frac{x}{2}} \tag{15}$$

$$(n = 1, 2, \ldots).$$

On peut s'en assurer en multipliant et en divisant les sommes partielles des séries considérées par $2\sin\frac{x}{2}$ et en procédant aux transformations trigonométriques nécessaires au numérateur. Les fonctions (15) sont majorées dans leur ensemble sur $[\varepsilon, 2\pi - \varepsilon]$:

$$|D_n(x)| \leqslant \frac{1}{2\sin(\varepsilon/2)}, \quad |K_n(x)| \leqslant \frac{1}{\sin(\varepsilon/2)} \quad (n = 1, 2, \ldots);$$

de plus $n^{-\alpha} \geqslant (n+1)^{-\alpha}$ et $n^{-\alpha} \to 0$, donc d'après le critère de Dirichlet les séries (14) convergent uniformément sur $[\varepsilon, 2\pi - \varepsilon]$.

## § 9. Intégration et dérivation des séries uniformément convergentes

THÉORÈME 1. *Soit donnée une suite $\{f_n(x)\}$ de fonctions (à valeurs complexes) convergeant vers une fonction $f$ sur $[a, b]$. Si la convergence est uniforme sur $[a, b]$, alors*

$$\lim_{n\to\infty} \int_a^x f_n(t)\,dt = \int_a^x f(t)\,dt \tag{1}$$

*uniformément sur* $[a, b]$. *En particulier* (*pour* $x = b$),

$$\lim_{n\to\infty} \int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt. \tag{2}$$

DÉMONSTRATION. De l'hypothèse du théorème il s'ensuit (cf. § 8, théorème 2) que la fonction limite $f$ est continue sur $[a, b]$ et

$$\max_{a\leqslant t\leqslant b} |f_n(t) - f(t)| = r_n \to 0 \quad (n \to \infty).$$

Donc

$$\left| \int_a^x f_n(t)\,dt - \int_a^x f(t)\,dt \right| \leqslant \int_a^x |f_n(t) - f(t)|\,dt \leqslant \int_a^b r_n\,dt = (b-a)\,r_n,$$

où le second membre ne dépend pas de $x$ et tend vers 0 pour $n \to \infty$, ce qui prouve le théorème.

THÉORÈME 2. *Toute série*

$$S(x) = u_0(x) + u_1(x) + u_2(x) + \ldots \tag{3}$$

*de fonctions* (*à valeurs complexes*) *uniformément convergente sur* $[a, b]$ *peut être intégrée terme à terme* ($a \leqslant x_0 \leqslant b$):

$$\int_{x_0}^x S(t)\,dt = \int_{x_0}^x u_0(t)\,dt + \int_{x_0}^x u_1(t)\,dt + \ldots \tag{4}$$

*La série* (4) *ainsi obtenue converge uniformément sur* $[a, b]$.
*En particulier*

$$\int_a^b S(t)\,dt = \int_a^b u_0(t)\,dt + \int_a^b u_1(t)\,dt + \ldots \tag{5}$$

DÉMONSTRATION. On remarquera que $S(x)$ est une fonction continue sur $[a, b]$, car elle est la somme d'une série uniformément convergente de fonctions continues sur $[a, b]$. Soit

$$S_n(x) = \sum_0^n u_k(x).$$

Comme la série (3) converge uniformément vers $S(x)$, alors

$$\sup_{x\in[a,\,b]} |S_n(x) - S(x)| = r_n \to 0 \quad (n \to \infty).$$

Donc

$$\left|\int_{x_0}^{x} S(t)\,dt - \sum_{0}^{n}\int_{x_0}^{x} u_k(t)\,dt\right| = \left|\int_{x_0}^{x} S(t)\,dt - \int_{x_0}^{x} S_n(t)\,dt\right| =$$

$$= \left|\int_{x_0}^{x} [S(t) - S_n(t)]\,dt\right| \leqslant \int_{a}^{b} |S(t) - S_n(t)|\,dt \leqslant$$

$$\leqslant (b-a)\,r_n \to 0 \quad (n \to \infty),$$

c. q. f. d.

THÉORÈME 3. *Soit donnée sur un intervalle* $[a, b]$ *une série*

$$u_0(x) + u_1(x) + u_2(x) + \ldots \tag{6}$$

*de fonctions (à valeurs complexes) possédant une dérivée continue.*

*Si la série* (6) *converge en un point* $x_0 \in [a, b]$ *et si la série dérivée*

$$u_0'(x) + u_1'(x) + u_2'(x) + \ldots \tag{7}$$

*converge uniformément sur* $[a, b]$, *alors la série* (6) *converge uniformément sur* $[a, b]$ *et la dérivée de sa somme* $S(x)$ *est égale à la somme de la série* (7).

Ainsi donc,

$$S(x) = u_0(x) + u_1(x) + u_2(x) + \ldots, \tag{8}$$

$$S'(x) = u_0'(x) + u_1'(x) + u_2'(x) + \ldots, \quad x \in [a, b]. \tag{9}$$

DÉMONSTRATION. Par hypothèse, la série (7) est uniformément convergente sur $[a, b]$ et ses termes sont des fonctions continues sur $[a, b]$. Donc, sa somme que nous désignons provisoirement par $\varphi(x)$ est une fonction continue sur $[a, b]$. Le théorème 2 nous dit que la série (7) donne par intégration une série uniformément convergente sur $[a, b]$:

$$\int_{x_0}^{x} \varphi(t)\,dt = \int_{x_0}^{x} u_0'(t)\,dt + \int_{x_0}^{x} u_1'(t)\,dt + \ldots \quad (a \leqslant x \leqslant b).$$

En appliquant le théorème de Newton-Leibniz, on aura

$$\int_{x_0}^{x} \varphi(t)\,dt = \sum_{0}^{\infty} [u_k(x) - u_k(x_0)]. \tag{10}$$

La série de fonctions entre crochets du second membre de (10) converge uniformément sur $[a, b]$. La série $\sum_{0}^{\infty} u_k(x_0)$ converge par

hypothèse et, comme ses termes sont constants, on peut la traiter comme une série uniformément convergente sur $[a, b]$. Donc, la série $\sum_0^\infty u_k(x)$ converge uniformément en tant que somme de deux séries uniformément convergentes; soit $S(x)$ sa somme. La relation (10) s'écrit maintenant:

$$S(x) = S(x_0) + \int_{x_0}^{x} \varphi(t)\, dt.$$

Or, $S'(x) = \varphi(x)$. Ceci achève la démonstration du théorème.

EXEMPLE 1. La série

$$S(x) = \frac{\cos x}{1^\alpha} + \frac{\cos 2x}{2^\alpha} + \frac{\cos 3x}{3^\alpha} + \dots \tag{11}$$

converge uniformément pour $\alpha > 1$ sur l'axe réel tout entier d'après le critère de Weierstrass, car

$$|n^{-\alpha} \cos nx| \leqslant n^{-\alpha}, \quad \forall x \in ]-\infty, \infty[$$

et

$$\sum_1^\infty n^{-\alpha} < \infty \quad (\alpha > 1).$$

Une dérivation formelle de la série (11) nous donne

$$\varphi(x) = \frac{-\sin x}{1^{\alpha-1}} - \frac{\sin 2x}{2^{\alpha-1}} - \frac{\sin 3x}{3^{\alpha-1}} - \dots \tag{12}$$

Cette série converge uniformément sur $]-\infty, \infty[$ pour $\alpha > 2$. Donc, pour $\alpha > 2$

$$S'(x) = \varphi(x). \tag{13}$$

Traitons le cas $1 < \alpha \leqslant 2$. Le critère de Weierstrass ne passe pas pour la série (12). Cependant, la série (12) converge uniformément sur l'intervalle $[\varepsilon, 2\pi - \varepsilon]$ pour tout $\varepsilon > 0$ (cf. § 8, exemple 4). La série (11) étant convergente sur cet intervalle, le théorème 3 affirme la validité de (13) sur l'intervalle $[\varepsilon, 2\pi - \varepsilon]$, quelque petit que soit $\varepsilon > 0$, donc sur l'intervalle $]0, 2\pi[$.

Comme les termes de la série (11) sont $2\pi$-périodiques, on vient de prouver que pour $1 < \alpha \leqslant 2$, la série (11) peut être dérivée terme à terme pour toutes les valeurs de $x \in ]-\infty, \infty[$ à l'exception des points $x_k = 2k\pi$ $(k = 0, \pm 1, \pm 2, \dots)$.

EXEMPLE 2. Soit donnée une fonction $f_n(x)$ continue sur $[0, 1]$, linéaire sur chacun des intervalles $[0, 1/2n]$ et $[1/2n, 1/n]$ et telle que $f_n(0) = f(1/n) = 0$, $f_n(1/2n) = \alpha_n$, $f_n(x) \equiv 0$ sur $[1/n, 1]$, où $\alpha_n$

est une suite quelconque de nombres (fig. 107). Il est alors évident que $\lim\limits_{n\to\infty} f_n(x) = 0$ pour tous les $x \in [0, 1]$ et

$$\int_0^1 f_n(x)\,dx = \int_0^{1/2n} 2n\alpha_n x\,dx + \int_{1/2n}^{1/n} 2n\alpha_n\left(\frac{1}{n}-x\right)dx = \frac{\alpha_n}{2n}.$$

Il est évident d'autre part que

$$r_n = \sup_{x\in[0,\,1]} |f_n(x) - 0| = \alpha_n,$$

donc la suite $\{f_n(x)\}$ converge uniformément si et seulement si $\alpha_n \to 0$. L'égalité

$$\int_0^1 f_n(x)\,dx \to \int_0^1 f(x)\,dx \quad (f(x)\equiv 0) \quad (14)$$

a lieu si et seulement si $\alpha_n/2n \to 0$ $(n\to\infty)$.

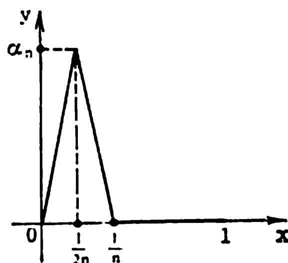


Fig. 107

On voit que la convergence uniforme de $f_n$ vers $f = 0$ sur $[0, 1]$ (c'est-à-dire lorsque $\alpha_n \to 0$) entraîne celle des intégrales (14), ce qui est en accord avec le théorème 2. Mais la suite $\{f_n\}$ peut converger *non uniformément*, alors que la propriété (14) a lieu tout de même, par exemple pour $\alpha_n = 1$. Ceci montre que la convergence uniforme de la suite est une condition suffisante mais pas nécessaire de convergence de la suite des intégrales vers l'intégrale de la fonction limite. Enfin, pour $\alpha_n = n$, la suite $\{f_n\}$ converge non uniformément vers 0 tandis que la propriété (14) n'a pas lieu.

Donc, si la suite $\{f_n\}$ converge non uniformément, alors il est possible que la suite des intégrales $\int_a^b f_n(x)\,dx$ converge vers l'intégrale de la fonction limite $\int_a^b f(x)\,dx$ ou vers un autre nombre (vers 1/2 et non vers 0 pour $\alpha_n = n$) ou pas du tout.

EXEMPLE 3. De l'égalité $(1-z)^{-1} = 1 + z + z^2 + \ldots$ $(z = \rho e^{i\theta}$, $\rho < 1)$ il s'ensuit que

$$\frac{1+\rho e^{i\theta}}{2(1-\rho e^{i\theta})} = \frac{1}{2} + \rho e^{i\theta} + \rho^2 e^{i2\theta} + \ldots,$$

et en séparant les parties réelle et imaginaire,

$$P(\rho, \theta)=\frac{1}{2}\,\frac{1-\rho^2}{1-2\rho\cos\theta+\rho^2}=\frac{1}{2}+\rho\cos\theta+\rho^2\cos 2\theta+\ldots, \qquad (15)$$

$$Q(\rho, \theta)=\frac{\rho\sin\theta}{1-2\rho\cos\theta+\rho^2}=\rho\sin\theta+\rho^2\sin 2\theta+\ldots \qquad (16)$$

La fonction $P(\rho, \theta)$ s'appelle *noyau de Poisson* [1], $Q(\rho, \theta)$ est la *fonction conjuguée de* $P(\rho, \theta)$.

Ces fonctions sont harmoniques (pour $\rho < 1$), c'est-à-dire qu'elles sont solutions de l'équation différentielle de Laplace [2] en coordonnées polaires

$$\Delta u \equiv \frac{\partial^2 u}{\partial \rho^2}+\frac{1}{\rho}\frac{\partial u}{\partial \rho}+\frac{1}{\rho^2}\frac{\partial^2 u}{\partial \theta^2}=0. \qquad (17)$$

En effet, chaque terme de la série (15) est une fonction harmonique
$(\rho^n\cos n\theta)'_\rho = n\rho^{n-1}\cos n\theta$, $(\rho^n\cos n\theta)''_\rho = n(n-1)\rho^{n-2}\cos n\theta$,

$$(\rho^n\cos n\theta)''_\theta = -n^2\rho^n\cos n\theta,$$

$$\Delta(\rho^n\cos n\theta) = \rho^{n-2}\cos n\theta\,[n(n-1)+n-n^2] = 0.$$

De façon analogue, $\Delta(\rho^n\sin n\theta) = 0$.

La dérivation terme à terme des séries (15) et (16) est licite, car les séries (15) et (16) et les séries dérivées (une ou deux fois) convergent uniformément pour $\rho \leqslant \rho_0 < 1$, où $\rho_0$ est un nombre quelconque $<1$.

Signalons qu'une fonction $u(x, y)$, où $x$ et $y$ sont des coordonnées cartésiennes, est *harmonique dans un domaine* $\Omega$ de points $(x, y)$ si elle est solution de l'équation différentielle

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0.$$

Cette équation est de la forme (17) en coordonnées polaires.

## § 10. Produit de séries absolument convergentes

Soient les séries

$$S=\sum_{k=0}^{\infty} u_k, \quad \sigma=\sum_{l=0}^{\infty} v_l \qquad (1)$$

à termes réels ou complexes, uniformément convergentes. Numérotons les couples $(k, l)$, où $k = 0, 1, 2, \ldots$, $l = 0, 1, 2, \ldots$, de la manière suivante:

$$(k_1, l_1), (k_2, l_2), (k_3, l_3), \ldots, \qquad (2)$$

[1] Denis Poisson, mathématicien français (1781-1840).
[2] Pierre Simon Laplace, mathématicien et physicien français (1749-1827).

chaque couple ne figurant qu'une fois dans la suite. Montrons que

$$S\sigma = \sum_{i=0}^{\infty} u_{k_i} v_{l_i} \tag{3}$$

et que la série du second membre converge absolument.

Si donc on forme une série avec les produits $u_k v_l$, alors cette série convergera absolument vers $S\sigma$.

Pour prouver cette proposition, composons les séries des modules $|u_k|$ et $|v_l|$:

$$\bar{S} = \sum_{k=0}^{\infty} |u_k|, \quad \bar{\sigma} = \sum_{l=0}^{\infty} |v_k|. \tag{1'}$$

Posons

$$\bar{S}_N = \sum_{k=0}^{N} |u_k|, \quad \bar{\sigma}_N = \sum_{l=0}^{N} |v_l|.$$

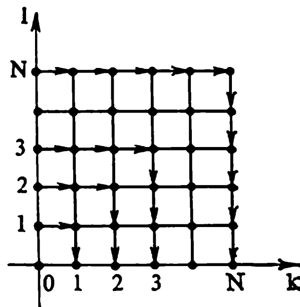


Fig. 108

Ordonnons les couples $(k, l)$ de la manière suivante (fig. 108) : (0, 0), (0, 1), (1, 1), (1, 0), (0, 2), (1, 2), (2, 2), (2, 1), (2, 0), . . . Alors

$$\bar{S}\bar{\sigma} = \lim_{N\to\infty} \bar{S}_N \lim_{N\to\infty} \bar{\sigma}_N = \lim_{N\to\infty} (\bar{S}_N \cdot \bar{\sigma}_N) =$$
$$= \lim_{N\to\infty} (|u_0|\cdot|v_0| + |u_0|\cdot|v_1| + |u_1|\cdot|v_1| + |u_1|\cdot|v_0| +$$
$$+ |u_0|\cdot|v_2| + |u_1|\cdot|v_2| + |u_2|\cdot|v_2| + |u_2|\cdot|v_1| +$$
$$+ |u_2|\cdot|v_0| + \ldots + |u_N|\cdot|v_0|). \tag{4'}$$

Ceci montre que la somme du second membre tend pour $N \to \infty$ vers une limite égale à $\bar{S}\bar{\sigma}$, et comme les termes de cette somme sont positifs, le nombre $\bar{S}\bar{\sigma}$ est égal à la somme de la série

$$\bar{S}\bar{\sigma} = |u_0|\cdot|v_0| + |u_0|\cdot|v_1| + |u_1|\cdot|v_1| + \ldots \tag{5'}$$

Les termes de cette série étant positifs, on peut les permuter sans modifier la convergence et la somme $\bar{S}\bar{\sigma}$.

Nous avons prouvé la proposition seulement pour les séries (1').

Supposons maintenant que

$$S_N = \sum_{k=0}^{N} u_k, \quad \sigma_N = \sum_{l=0}^{N} v_l.$$

Comme dans (4'), la convergence des séries (1) nous donne

$$S\sigma = \lim_{N\to\infty} (S_N \sigma_N) = \lim_{N\to\infty} (u_0v_0 + u_0v_1 + u_1v_1 + u_1v_0 + \ldots$$
$$\ldots + u_N v_0). \tag{4}$$

Donc, lorsque $N \to \infty$, le second membre de (4) tend vers une limite égale à $S\sigma$. Or, nous avons déjà prouvé que la série (5') était convergente. Ceci montre que la série

$$u_0v_0 + u_0v_1 + u_1v_1 + u_1v_0 + \ldots \tag{5}$$

converge absolument.

En vertu de (4), la somme de cette série est égale à $S\sigma$:

$$S\sigma = u_0v_0 + u_0v_1 + u_1v_1 + u_1v_0 + \ldots$$

Nous venons de prouver l'égalité (3) pour le procédé de numération (2) des couples $(k, l)$. Mais comme la série (5) est absolument convergente, l'égalité (3) est valable pour tout autre procédé de numération.

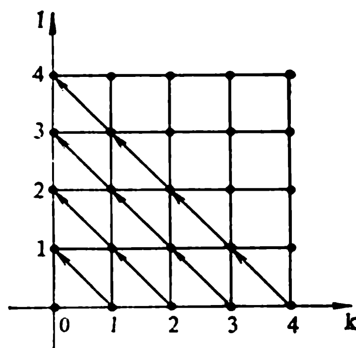


Fig. 109

EXEMPLE. La série

$$\psi(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots \tag{6}$$

converge absolument pour toute valeur complexe $z$ (on dit encore qu'elle converge absolument sur le plan complexe tout entier). On s'en assure en appliquant le critère de D'Alembert à la série de terme général $|z|^n/n!$

Pour deux nombres complexes $u$ et $v$ quelconques on a (voir explications plus bas)

$$\psi(u)\psi(v) = \left(1 + u + \frac{u^2}{2!} + \ldots\right)\left(1 + v + \frac{v^2}{2!} + \ldots\right) =$$

$$= 1 + u + v + \frac{u^2}{2!}1 + u\cdot v + 1\frac{v^2}{2!} + \frac{u^3}{3!}\cdot 1 + \frac{u^2}{2!}v + u\frac{v^2}{2!} + 1\cdot\frac{v^3}{3!} +$$

$$+ \ldots = 1 + (u+v) + \frac{1}{2!}(u^2 + 2uv + v^2) +$$

$$+ \frac{1}{3!}(u^3 + 3u^2v + 3uv^2 + v^3) + \ldots =$$

$$= 1 + (u+v) + \frac{(u+v)^2}{2!} + \frac{(u+v)^3}{3!} + \ldots = \psi(u+v). \tag{7}$$

Dans la deuxième égalité nous avons disposé les produits $\frac{u^k}{k!}\frac{v^l}{l!}$ dans l'ordre indiqué sur la figure 109 et utilisé l'égalité (3) pour les séries absolument convergentes. La série ainsi obtenue est absolument convergente, ainsi qu'on l'a prouvé pour le cas général. Dans

les égalités suivantes, on a groupé certains termes entre parenthèses sans altérer la convergence.

Nous avons prouvé l'égalité importante

$$\psi(u+v) = \psi(u)\cdot\psi(v), \tag{8}$$

quels que soient $u$ et $v$ complexes. Nous reparlerons plus loin au § 13 de cette égalité.

## § 11. Séries entières

On appelle *série entière* une série de la forme

$$a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \ldots, \tag{1}$$

où $a_k$ sont des constantes et $z$ une variable. Les nombres $a_k$ et la variable $z$ peuvent être complexes. Dans la suite, $z$ désignera un nombre complexe (un point du plan complexe), et $x$, un nombre réel (un point de la droite réelle).

Le théorème suivant est fondamental en théorie des séries entières.

THÉORÈME 1 (FONDAMENTAL). *Pour la série entière* (1) *il existe un nombre positif $R$ fini ou infini* ($0 \leqslant R \leqslant \infty$) *doué des propriétés suivantes*:

1) *La série converge absolument dans un disque ouvert $|z| < R$ du plan complexe et diverge aux points $z$ tels que $|z| > R$.*

2) *Le nombre $R$ est défini par la formule*

$$R = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}. \tag{2}$$

On convient que $\frac{1}{0} = \infty$ et $\frac{1}{\infty} = 0$, de sorte que $R = 0$ ou $\infty$ selon que la limite supérieure est infinie ou nulle.

Le disque ouvert $|z| < R$ s'appelle *disque de convergence de la série entière*. Pour $R = \infty$ c'est le plan complexe tout entier. Pour $R = 0$ la série entière ne possède qu'un point de convergence, en l'occurrence le point $z = 0$; $R$ s'appelle *rayon de convergence de la série* (1).

REMARQUE 1. Il n'existe visiblement qu'un seul nombre $R$ vérifiant la proposition 1) du théorème 1.

REMARQUE 2. Si la limite $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}$ existe, alors elle est égale à $\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$. Donc

$$R = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}.$$

Le lecteur non initié à la notion de limite supérieure peut néanmoins suivre la démonstration du théorème en admettant que la limite indiquée existe pour la série entière considérée. Dans ce cas, il faut partout remplacer lim. sup par lim.

DÉMONSTRATION DU THÉORÈME 1. Supposons que $R$ est donné par la formule (2). La série converge au point $z = 0$. On admettra d'autre part que $|z| > 0$. Considérons en outre la série des modules des termes de la série (1):

$$|a_0| + |a_1 z| + |a_2 z^2| + \dots \qquad (1')$$

Désignons le terme général de (1′) par

$$u_n = |a_n z^n| \quad (n = 0, 1, 2, \dots). \qquad (3)$$

Le critère de Cauchy (cf. § 4, théorème 3, c)) nous dit que la série (1′)

$$\text{converge si } \limsup_{n\to\infty} \sqrt[n]{u_n} < 1,$$

$$\text{diverge si } \limsup_{n\to\infty} \sqrt[n]{u_n} > 1,$$

la variable $\sqrt[n]{u_n}$ étant illimitée dans ce cas, mais

$$\limsup \sqrt[n]{u_n} = \limsup \sqrt[n]{|a_n z^n|} = \limsup \left(|z| \sqrt[n]{|a_n|}\right) =$$

$$= |z| \limsup \sqrt[n]{|a_n|} = \frac{|z|}{R}.$$

Nous avons sorti le nombre fini $|z| > 0$ du signe de la limite supérieure.

De ce qui précède il suit que si $|z| < R$, c'est-à-dire que $|z|/R < 1$, alors la série (1′) converge, donc la série (1) converge absolument; si $|z| > R$, c'est-à-dire que $|z|/R > 1$, alors la série (1′) diverge et son terme général $|a_n z^n|$ est illimité, donc le terme général $a_n z^n$ de la série (1) ne tend pas vers 0 pour $n \to \infty$ et de ce fait n'est pas justiciable de la condition nécessaire (cf. § 1). Ceci exprime que la série (1) diverge.

Nous avons ainsi prouvé que le nombre $R$ défini par (2) jouit de la propriété suivante:

si $|z| < R$, alors la série (1) converge absolument,
si $|z| > R$, la série (1) diverge.

Ceci achève la démonstration du théorème fondamental.

Dans la suite on désignera par $\sigma_q$ le disque fermé $|z| \leqslant q$ du plan complexe.

A noter que la série entière converge sur le disque ouvert $|z| < R$ en général non uniformément. Cependant on a le théorème suivant.

THÉORÈME 2. *La série entière* (1) *converge absolument et uniformément sur tout disque* $\sigma_q = \{z : |z| \leqslant q\}$, *où* $q < R$ *et* $R$ *est le rayon de convergence de la série* (1).

DÉMONSTRATION. En effet, supposons que $q < R$, $q$ est alors réel, c'est-à-dire est un point de l'axe $Ox$ appartenant au disque ouvert de convergence de la série (1). Donc, la série entière converge absolument en ce point, autrement dit

$$\sum_{n=0}^{\infty} |a_n q^n| < \infty.$$

D'autre part, pour tout point $z \in \sigma_q$ on a

$$|a_n z^n| \leqslant |a_n q^n| \quad (n = 0, 1, 2, \ldots).$$

Comme les seconds membres de ces inégalités ne dépendent pas de $z \in \sigma_q$ et que la série de terme général $|a_n q^n|$ converge, le critère de Weierstrass (cf. § 8, théorème 1) nous dit que la série (1) converge absolument et uniformément sur $\sigma_q$.

THÉORÈME 3. *La somme*

$$S(z) = a_0 + a_1 z + a_2 z^2 + \ldots$$

*de la série entière* (1) *est une fonction continue sur le disque de convergence* $|z| < R$ *de cette série.*

En effet, les termes de la série sont des fonctions continues de $z$ et la série converge uniformément sur le disque $\sigma_q$, $q < R$. Donc, d'après un théorème connu de la théorie des séries uniformément convergentes (cf. § 8, théorème 2) la somme $S(z)$ de la série est une fonction continue sur $\sigma_q$, donc sur le disque $|z| < R$ tout entier, puisque $q$ est arbitraire.

Pour calculer le rayon de convergence de la série entière, on dispose de la formule (2), mais souvent il est plus commode d'utiliser le critère de D'Alembert.

Supposons qu'éxiste la limite (finie ou infinie)

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \tag{4}$$

que nous désignerons provisoirement par $1/R_1$. Alors (cf. (3))

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim_{n\to\infty} \frac{|a_{n+1} z^{n+1}|}{|a_n z^n|} = |z| \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{|z|}{R_1}.$$

En vertu du critère de D'Alembert (§ 4, théorème 2), si $|z| < R_1$, alors la série (1'), donc la série (1), converge ; si $|z| > R_1$, alors $|u_n| \to \infty$ et la série (1) diverge. Or, il n'existe qu'un seul nombre doué de telles propriétés, donc $R_1 = R$ (cf. théorème 1).

Nous avons ainsi montré que, si la limite (4) existe, elle est égale à $1/R$:

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=\frac{1}{R}, \tag{5}$$

où $R$ est le rayon de convergence de la série entière (1).

A noter que nous venons de montrer indirectement que si la limite (4) (finie ou infinie) existe, alors elle est égale à lim. sup $\sqrt[n]{|a_n|}$.

Remarque 3. L'exposé de la théorie des séries entières commence généralement par le théorème d'Abel suivant.

Théorème d'Abel. *Si la série entière* (1) *converge en un point* $z_0 \neq 0$ *du plan complexe, alors elle converge absolument et uniformément dans le disque fermé* $|z| \leqslant q$, *où* $q$ *est un nombre quelconque vérifiant la double inégalité* $0 < q < |z_0|$.

Démonstration. Ce théorème est maintenant un corollaire des théorèmes 1 et 2. En effet, le point $z_0$ étant un point de convergence de la série (1), $|z_0|$ ne peut être strictement supérieur à $R$. Donc $|z_0| \leqslant R$, $0 < q < |z_0| \leqslant R$ et $q < R$. Le théorème 2 nous dit alors que la série entière (1) converge absolument et uniformément sur $|z| \leqslant q$.

Exemples.

$$1+z+z^2+\ldots, \tag{6}$$

$$1+\frac{z}{1^\alpha}+\frac{z^2}{2^\alpha}+\frac{z^3}{3^\alpha}+\ldots \quad (\alpha>0), \tag{7}$$

$$1+z+2!\,z^2+3!\,z^3+\ldots \tag{8}$$

Les formules (2) ou (5) nous permettent de conclure que le rayon de convergence des séries (6) et (7) est égal à 1 et celui de la série (8), à 0.

La somme de la série (6) (qui est une progression géométrique) est égale à $(1-z)^{-1}$ dans le disque ouvert $|z| < 1$ et le reste

$$r_n(z)=\sum_{n+1}^{\infty} z^k=\frac{z^{n+1}}{1-z}\to 0 \quad (n\to\infty).$$

Cependant, la convergence n'est pas uniforme dans ce disque. Ceci a lieu déjà pour les $z = x$ sur l'intervalle $0 < x < 1$, l'inégalité

$$\varepsilon > \frac{x^{n+1}}{1-x} \tag{9}$$

ne pouvant être vérifiée pour tous les $x$ indiqués, quel que soit $n$.

En effet, si $x$ est très proche de 1, le numérateur, lui aussi, sera proche de 1 et le dénominateur, de 0. Donc, la fraction peut être rendue plus grande que $\varepsilon$.

La série (7) converge uniformément pour $\alpha > 1$ sur le disque fermé $|z| \leqslant 1$, puisque pour $|z| \leqslant 1$

$$|z^\alpha k^{-\alpha}| \leqslant k^{-\alpha} \text{ et } \sum k^{-\alpha} < \infty.$$

Si $\alpha = 1$, la série (7) diverge au point $z = 1$ situé sur le cercle de convergence.

La série (8) ne converge qu'au point $z = 0$.

## § 12. Dérivation et intégration des séries entières

THEOREME 1. *La série entière*

$$a_0 + a_1 z + a_2 z^2 + \ldots \tag{1}$$

*et la série dérivée*

$$a_1 + 2a_2 z + 3a_3 z^2 + \ldots \tag{2}$$

*ont même rayon de convergence.*

REMARQUE. La définition de la continuité et de la dérivée d'une fonction à variable complexe $f(z)$ est la même que pour une fonction à variable réelle. La seule différence c'est que le $\delta$-voisinage du point $z_0$ est un disque ouvert de rayon $\delta$ et de centre $z_0$. En partant de cette définition, on établit que la dérivée de la fonction puissance $z^n$ a pour expression $(z^n)' = nz^{n-1}$.

DEMONSTRATION DU THEOREME 1. On admettra que $R$ est le rayon de convergence de la série (1) et $R'$ celui de la série (2). Prouvons ce théorème sous l'hypothèse qu'existe la limite (finie ou infinie)

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = \frac{1}{R}. \tag{3}$$

On a

$$\frac{1}{R'} = \lim_{n\to\infty} \sqrt[n]{|(n+1)\,a_{n+1}|} =$$
$$= \lim_{n\to\infty} \sqrt[n]{n+1} \cdot \lim_{n\to\infty} \sqrt[n]{|a_{n+1}|} = 1 \cdot \frac{1}{R} = \frac{1}{R},$$

donc $R = R'$.

Dans le cas général où la limite (3) n'existe pas, on a

$$\lim_{n\to\infty}. \sup \sqrt[n]{|a_n|} = \frac{1}{R}$$

et alors

$$\frac{1}{R'} = \lim_{n\to\infty}. \sup \sqrt[n]{n\,|a_n|} = \lim_{n\to\infty} \sqrt[n]{n} \lim_{n\to\infty}. \sup \sqrt[n]{|a_n|} = 1 \cdot \frac{1}{R} = \frac{1}{R}.$$

La deuxième égalité doit être justifiée, c'est-à-dire qu'il nous faut prouver que si $\alpha_n$, $\beta_n > 0$ et $\alpha_n \to 1$, alors

$$\lim.\sup(\alpha_n\beta_n) = \lim.\sup \beta_n. \tag{4}$$

Il existe en effet une suite extraite $\{n_k\}$ telle que

$$\lim.\sup \beta_n = \lim \beta_{n_k} = \lim \alpha_{n_k} \cdot \lim \beta_{n_k} = \lim (\alpha_{n_k}\beta_{n_k}) \leqslant \lim.\sup(\alpha_n\beta_n). \tag{5}$$

Il existe aussi une suite extraite $\{n_k\}$ telle que

$$\lim.\sup(\alpha_n\beta_n) = \lim(\alpha_{n_k}\beta_{n_k}) = \frac{\lim(\alpha_{n_k}\beta_{n_k})}{\lim \alpha_{n_k}} =$$

$$= \lim \beta_{n_k} \leqslant \lim.\sup \beta_n. \tag{6}$$

De (5) et de (6) on déduit (4).

THÉORÈME 2. *La série entière*

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots \quad (|z| < R) \tag{7}$$

*peut être dérivée dans les limites de son disque de convergence* $|z| < R$, *autrement dit*

$$f'(z) = a_1 + 2a_2 z + 3a_3 z^2 + \ldots \quad (|z| < R). \tag{8}$$

DÉMONSTRATION. On prouvera ce théorème uniquement sous l'hypothèse que $z = x$ est une variable réelle, ce qui nous permettra de ramener le problème à un fait bien connu de la théorie des séries réelles.

La série entière (7) s'écrit maintenant

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots, \quad x \in ]-R, R[. \tag{7'}$$

Cette série possède maintenant non plus un disque mais un intervalle de convergence $]-R, R[$. La série dérivée s'écrit:

$$\varphi(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots \tag{8'}$$

Soit $\varphi(x)$ la somme de cette série. Le théorème précédent nous dit qu'elle converge sur l'intervalle $]-R, R[$. Nous savons que les deux séries convergent uniformément sur l'intervalle $[-q, q]$, où $q < R$. En outre, les termes de la deuxième série sont continus et sont les dérivées de ceux de la première. Donc, en vertu de théorème 3 du § 9 de la théorie des séries uniformément convergentes, on a

$$\varphi(x) = f'(x) \tag{9}$$

sur l'intervalle $[-q, q]$ et par suite sur l'intervalle $]-R, R[$, car $q$ est arbitraire.

Signalons qu'en vertu du théorème 2, la série (1) peut être dérivée terme à terme autant de fois qu'on le veut. Une dérivation d'ordre $k$ nous donne

$$f^{(k)}(z) = k!a_k + (k+1)k \ldots 2a_{k+1}z + \ldots$$

pour tous les $|z| < R$. Si l'on pose $z = 0$, on obtient

$$f^{(k)}(0) = k!a_k$$

ou

$$a_k = \frac{f^{(k)}(0)}{k!} \quad (k = 0, 1, 2, \ldots).$$

De là il suit en particulier que *le développement de la fonction* $f(z)$ *en série entière* (cf. (1)) *dans un disque* $|z| < R$ (*ou dans l'intervalle* $(-R < x < R)$ *dans le cas d'une fonction* $f(x)$ *à variable réelle*) *est unique.*

Donc, la somme $f(z)$ de la série entière (7) de rayon de convergence $R > 0$ peut se mettre sous la forme suivante:

$$f(z) = f(0) + \frac{f'(0)}{1!}z + \frac{f''(0)}{2!}z^2 + \ldots = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} z^k. \qquad (10)$$

La série du second membre s'appelle *série de Taylor de la fonction* $f(z)$ *suivant les puissances de* $z$.

Par conséquent, *si la série entière* (1) *possède un rayon de convergence* $R > 0$, *alors elle est la série de Taylor de sa somme* $f(z)$.

L'intégration terme à terme des séries entières impliquerait l'introduction de l'intégrale curviligne d'une fonction à variable complexe. Aussi se contentera-t-on d'étudier ce problème seulement pour les séries entières

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots \qquad (11)$$

d'une variable réelle $x$ $(z = x)$.

Considérons une série entière (11) d'intervalle de convergence $]-R, R[$, où $R \in ]0, \infty]$. Les nombres $a_k$ peuvent être réels ou complexes. Soient $x_0$ un point fixe de $]-R, R[$, $x$ un point variable de $]-R, R[$. Choisissons $q > 0$ de telle sorte que

$$-R < -q < x_0, \quad x < q < R.$$

La série entière (11) converge uniformément sur l'intervalle $[-q, q]$ qui est intérieur à l'intervalle de convergence de la série. Donc, cette série peut être intégrée terme à terme (cf. § 9, théorème 2) entre $x_0$ et $x$:

$$\int_{x_0}^{x} f(t)\,dt = a_0(x - x_0) + \frac{a_1}{2}(x^2 - x_0^2) + \frac{a_2}{3}(x^3 - x_0^3) + \ldots, \qquad (12)$$

$$-R < x, \ x_0 < R.$$

En particulier, pour $x_0 = 0$, on obtient

$$\int_0^x f(t)\,dt = a_0x + \frac{a_1}{2}x^2 + \frac{a_2}{3}x^3 + \ldots, \qquad -R < x < R. \tag{13}$$

EXEMPLE 1. Il est évident que

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - t^6 + \ldots$$

Cette série converge sur l'intervalle $]-1, 1[$ ($R = 1$). Donc, si $x \in ]-1, 1[$, on peut intégrer cette série entre 0 et $x$ (cette série converge uniformément sur tout intervalle fermé contenu dans l'intervalle de convergence):

$$\int_0^x \frac{dt}{1+t^2} = \text{Arctg}\, x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots \quad (-1 < x < 1).$$

La série obtenue converge aussi pour $x = 1$ (en tant que série de Leibniz). On démontre qu'elle converge vers Arctg $1 = \pi/4$, i. e.

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

EXEMPLE 2. La série de Taylor de la fonction $e^{-t^2}$ s'écrit (cf. chap. 4, § 16)

$$e^{-t^2} = 1 - t^2 + \frac{t^4}{2!} - \frac{t^6}{3!} + \ldots$$

Ici $R = \infty$. Donc, cette série s'intègre terme à terme:

$$\int_0^x e^{-t^2}\,dt = x - \frac{x^3}{3} + \frac{x^5}{5\cdot 2!} - \frac{x^7}{7\cdot 3!} + \ldots,$$

et nous obtenons l'expression de l'intégrale de Poisson par une série entière.

EXEMPLE 3. La série de Taylor de la fonction $y = \sin x$ s'écrit (cf. chap. 4, § 16)

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$$

Elle converge sur l'axe réel tout entier. Pour $x \neq 0$, on obtient

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \ldots \tag{14}$$

Si l'on admet que $\left.\frac{\sin x}{x}\right|_{x=0} = 1$, on trouve que (14) est vraie également pour $x = 0$. La série (14) converge uniformément sur

tout intervalle fini de l'axe des réels. En intégrant cette série, on obtient

$$\int_0^x \frac{\sin t}{t}\, dt = x - \frac{x^3}{3\cdot 3!} + \frac{x^5}{5\cdot 5!} - \frac{x^7}{7\cdot 7!} + \ldots$$

EXEMPLE 4. La série de Taylor de la fonction $y = \cos x^2$ est de la forme (cf. chap. 4, § 16)

$$\cos x^2 = 1 - \frac{x^4}{2!} + \frac{x^8}{4!} - \frac{x^{12}}{6!} + \ldots$$

Elle converge sur $]-\infty, \infty[$. En intégrant cette série, on obtient l'intégrale de Fresnel

$$\int_0^x \cos t^2\, dt = x - \frac{x^5}{5\cdot 2!} + \frac{x^9}{9\cdot 4!} - \frac{x^{13}}{13\cdot 6!} + \ldots$$

EXEMPLE 5. Comme

$$(\operatorname{sh} x)^{(n)} = \begin{cases} \operatorname{sh} x & \text{pour } n = 2k, \\ \operatorname{ch} x & \text{pour } n = 2k+1, \end{cases}$$

on a

$$(\operatorname{sh} x)^{(n)}\big|_{x=0} = \begin{cases} 0 & \text{pour } n = 2k, \\ 1 & \text{pour } n = 2k+1, \end{cases}$$

et la suite de Taylor de $\operatorname{sh} x$ s'écrit:

$$\operatorname{sh} x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots \tag{15}$$

La série entière étant convergente sur l'axe des réels tout entier (grâce au critère de D'Alembert), on peut la dériver:

$$\operatorname{ch} x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots \tag{16}$$

(le second membre converge uniformément sur tout intervalle fini).

## § 13. Fonctions $e^z$, $\sin z$ et $\cos z$ d'une variable complexe

Les fonctions $e^x$, $\sin x$ et $\cos x$ de la variable réelle $x$ sont bien connues. Elles sont définies sur l'axe des réels tout entier.

Du § 16 du chap. 4 on sait qu'elles se développent en les séries entières suivantes:

$$\begin{cases} e^x = 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \ldots, \\ \sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \ldots, \\ \cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \ldots \end{cases} \tag{1}$$

Au § 3 du chap. 5 on a défini la fonction $e^{ix}$, où $x$ est une variable réelle, au moyen de la formule d'Euler

$$e^{ix} = \cos x + i \sin x \tag{2}$$

Dans (2), substituons à $\cos x$ et à $\sin x$ leurs développements en série. On obtient alors le développement de $e^{ix}$ suivant les puissances de $x$:

$$e^{ix} = \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots\right) =$$
$$= 1 + \frac{ix}{1!} + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \dots$$

La fonction $e^z$, où $z = x + iy$, se définit naturellement de la manière suivante:

$$e^z = e^{x+iy} = e^x e^{iy}.$$

D'où

$$e^z = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots\right)\left(1 + \frac{iy}{1!} + \frac{(iy)^2}{2!} + \dots\right) =$$
$$= 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots$$

La multiplication est licite (cf. § 10, (7)), car les séries sont absolument convergentes.

On obtient en définitive que la fonction $e^z$ de la variable complexe $z$ se développe en la série entière

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \tag{3}$$

convergeant vers $e^z$ sur le plan complexe tout entier.

La série (3) est la série de Taylor de la fonction $e^z$.

Le rayon de convergence de la série (3) est $R = \infty$. En outre, des propriétés générales des séries entières (cf. § 10) il s'ensuit que la série (3) converge absolument pour tout $z$ et de plus converge uniformément (vers $e^z$) sur le disque $|z| \leqslant q$, quelque grand que soit le nombre positif $q$.

Les fonctions $\cos z$ et $\sin z$ de la variable complexe $z$ se définissent naturellement comme les sommes des séries entières suivantes

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots,$$
$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots$$

Le rayon de convergence de ces deux séries est $R = \infty$, donc les fonctions correspondantes sont définies pour tout $z$ complexe.

On vérifie immédiatement par comparaison des séries entières respectives que

$$\cos z = \frac{e^{iz}+e^{-iz}}{2}, \quad \sin z = \frac{e^{iz}-e^{-iz}}{2i} \tag{4}$$

pour tout $z$ complexe.

Les propriétés de la fonction exponentielle $e^u$ ($u$ complexe) nous donnent :

$$\cos(u+v) = \cos u \cos v - \sin u \sin v,$$
$$\sin(u+v) = \sin u \cos v + \cos u \sin v$$

pour tous $u$ et $v$ complexes.

Ces formules généralisent donc les formules classiques de trigonométrie pour $u$ et $v$ réels. Signalons que les fonctions $\sin z$ et $\cos z$ ne possèdent pas toutes les propriétés des fonctions réelles $\sin x$ et $\cos x$. En particulier, ces fonctions ne sont pas bornées sur le plan complexe.

En vertu de (4), pour $x$ réel on a

$$\cos ix = \frac{e^{-x}+e^{x}}{2} = \operatorname{ch} x \to \infty, \quad x \to \infty, \tag{5}$$

$$\sin ix = \frac{e^{-x}-e^{x}}{2i} = i \operatorname{sh} x \to \infty, \quad x \to \infty. \tag{6}$$

Les formules (5) et (6) établissent une relation entre la « trigonométrie complexe » et la « trigonométrie hyperbolique ».

La fonction $z = \ln w$ de la variable complexe $w$ se définit comme la réciproque de la fonction

$$w = e^z. \tag{7}$$

Si l'on écrit $w \neq 0$ sous la forme exponentielle

$$w = \rho e^{i\theta} \quad (\rho = |w| > 0),$$

alors l'équation (7) devient

$$\rho e^{i\theta} = e^x e^{iy} \quad (z = x + iy),$$

d'où

$$\rho = e^x, \quad \theta = y - 2k\pi,$$

c'est-à-dire que

$$x = \ln \rho, \quad y = \theta + 2k\pi \quad (k = 0, \pm 1, \pm 2, \ldots).$$

Donc,

$$z = \ln w = x + iy = \ln \rho + i(\theta + 2k\pi) =$$
$$= \ln |w| + i \arg w = \ln |w| + i \operatorname{Arg} w + i\, 2k\pi \tag{8}$$
$$(k = 0, \pm 1, \pm 2, \ldots),$$

où $\ln |w|$ ($|w| > 0$) est compris au sens ordinaire. On voit sur (8) que $\ln w$ ($w \neq 0$), de même d'ailleurs que $\arg w$, est une correspondance, que $w$ soit réelle ou complexe.

Par exemple, du point de vue de la théorie des fonctions d'une variable complexe, $\ln 1$ est égal à $2k\pi i$ ($k = 0, \pm 1, \pm 2, \ldots$). En analyse réelle, $\ln 1$ ne prend qu'une seule valeur, 0.

Sans nous appesantir sur la théorie des fonctions d'une variable complexe nous nous contenterons de faire encore une remarque au sujet de la formule

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots, \quad x \in ]-1, 1],$$

établie au § 16 du chap. 4 pour les $x$ réels. Si l'on remplace $x$ par $z$, avec

$$|z| < 1,$$

la série reste convergente. Sa somme est égale à $\ln(1+z)$, telle que nous l'avons définie plus haut.

Les fonctions d'une variable complexe développables en séries entières s'appellent *fonctions analytiques*. Elles font l'objet de la théorie des fonctions d'une variable complexe [1]).

Signalons enfin que si dans la série entière

$$a_0 + a_1 u + a_2 u^2 + \ldots$$

de disque de convergence $|u| < R$, on pose $u = z - z_0$, où $z_0$ est un nombre fixe (généralement complexe), on obtient la série entière

$$a_0 + a_1 (z - z_0) + a_2 (z - z_0)^2 + \ldots$$

appelée *série entière* de $z - z_0$.

Elle converge dans le disque (de convergence) $|z - z_0| < R$ et diverge pour $|z - z_0| > R$.

## § 14. Séries dans les calculs approchés

Dans ce paragraphe on se penchera sur le calcul approché des valeurs de fonctions élémentaires.

Commençons par le polynôme

$$P_n(x) = a_0 + a_1 x + \ldots + a_n x^n.$$

Le calcul de cette fonction pour $x = x_0$ se ramène à la réalisation d'un nombre fini d'additions et de multiplications. Cette fonction peut être calculée en tout point $x_0$ avec n'importe quelle précision.

Les autres fonctions élémentaires, telles $\sin x$, $\operatorname{Arctg} x$, ... se développent en une série entière.

---

[1]) Voir tome 2.

L'erreur commise en remplaçant une fonction par sa série de Taylor peut être déterminée à l'aide du reste.

Considérons la série entière

$$f(x) = c_0 + c_1x + c_2x^2 + \ldots, \quad x \in ]-R, R[, \tag{1}$$

d'intervalle de convergence $]-R, R[$. Cette série converge vers $f(x)$ à l'intérieur de l'intervalle de convergence à la vitesse d'une progression géométrique décroissante.

En effet, soient $q_1$ et $q$ des nombres arbitraires tels que $0 < q_1 < < q < R$. Alors la série (1) converge au point $x = q$ et ses termes forment une suite bornée ($|c_nq^n| \leqslant M$, $\forall n$). Donc, pour tous les $x \in [-q_1, q_1]$

$$|c_nx^n| = \left|c_nq^n\left(\frac{x}{q}\right)^n\right| \leqslant M\left(\frac{q_1}{q}\right)^n,$$

où $q_1/q < 1$.

On voit qu'il est plus favorable d'utiliser une série entière pour le calcul des valeurs de $f(x)$ aux points intérieurs à l'intervalle de convergence.

Si le point $x$ est confondu avec l'une des extrémités de l'intervalle $]-R, R[$, alors, même si la série converge en ce point, elle convergera moins vite qu'une progression géométrique décroissante. La vitesse de convergence est en général si lente qu'on n'a aucun intérêt à se servir de la série (1) pour calculer $f(x)$ en ce point. Ces faits seront illustrés plus bas sur des exemples concrets.

Commençons par le calcul du nombre $\pi$.

Dans l'exemple 1 du § 12, on a montré que

$$\operatorname{Arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots \quad x \in ]-1, 1[. \tag{2}$$

Cette série converge aussi au point $x = 1$. Montrons qu'elle converge vers $\operatorname{Arctg} 1 = \pi/4$. Ce fait n'a pas été prouvé au § 12.

Considérons l'identité

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - \ldots + (-1)^n x^{2n} + (-1)^{n+1}\frac{x^{2n+2}}{1+x^2}.$$

En intégrant cette identité sur $[0, 1]$, on trouve

$$\int_0^1 \frac{dx}{1+x^2} = \operatorname{Arctg} 1 = \frac{\pi}{4} = \int_0^1 dx - \int_0^1 x^2\,dx + \ldots$$

$$\ldots + (-1)^n \int_0^1 x^{2n}\,dx + (-1)^{n+1}\int_0^1 \frac{x^{2n+2}}{1+x^2}\,dx =$$

$$= 1 - \frac{1}{3} + \frac{1}{5} - \ldots + (-1)^n\frac{1}{2n+1} + \alpha_n,$$

où

$$\alpha_n = (-1)^{n+1} \int_0^1 \frac{x^{2n+2}}{1+x^2}\, dx.$$

Il est immédiat de voir que

$$|\alpha_n| \leqslant \int_0^1 x^{2n+2}\, dx = \frac{1}{2n+3} \to 0, \quad n \to \infty.$$

D'où

$$\left| \operatorname{Arctg} 1 - \sum_{k=0}^{n} \frac{(-1)^k}{2k+1} \right| \to 0, \quad n \to \infty,$$

c'est-à-dire que Arctg 1 est la somme de la série:

$$\operatorname{Arctg} 1 = \frac{\pi}{4} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$$

ou

$$\pi = 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}. \tag{3}$$

On remarque que cette série converge moins vite que toute progression géométrique décroissante.

Pour calculer $\pi$ à l'aide de la série (3) à $10^{-6}$ près, il faut prendre assez de termes pour que le reste soit inférieur à $10^{-6}$. Comme la série (3) est une série de Leibniz, alors son reste est inférieur au module du premier terme:

$$|R_n| = \left| 4 \sum_{k=n+1}^{\infty} \frac{(-1)^k}{2k+1} \right| < \frac{4}{2n+3}.$$

De là on voit que pour $n = 2 \cdot 10^6$, on a $|R_n| < 10^{-6}$. Donc, pour obtenir la précision voulue il faut deux millions de termes de la série (3).

A la main ce calcul est insensé. Avec une calculatrice, il ne présente aucun intérêt.

Exhibons une série qui converge plus vite vers le nombre $\pi$. Considérons a cet effet un nombre $\alpha$ tel que

$$\operatorname{tg} \alpha = 1/5.$$

Alors

$$\operatorname{tg} 2\alpha = \frac{2 \operatorname{tg} \alpha}{1 - \operatorname{tg}^2 \alpha} = \frac{2/5}{1 - 1/25} = \frac{5}{12},$$

$$\operatorname{tg} 4\alpha = \frac{2 \operatorname{tg} 2\alpha}{1 - \operatorname{tg}^2 2\alpha} = \frac{120}{119},$$

$$\operatorname{tg}\left(4\alpha - \frac{\pi}{4}\right) = \frac{\operatorname{tg} 4\alpha - \operatorname{tg}(\pi/4)}{1 + \operatorname{tg} 4\alpha \cdot \operatorname{tg}(\pi/4)} = \frac{1}{239}.$$

D'où

$$4\alpha - \frac{\pi}{4} = \operatorname{Arctg}(1/239),$$

$$\pi = 16\alpha - 4 \operatorname{Arctg}(1/239) = 16 \operatorname{Arctg}(1/5) - 4 \operatorname{Arctg}(1/239).$$

En se servant de la série (2), on obtient

$$\pi = 16 \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)\, 5^{2k+1}} - 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)\, 239^{2k+1}}.$$

Les deux dernières séries convergent assez vite (plus vite qu'une progression géométrique décroissante).

Il est immédiat de vérifier que le reste de la première série est inférieure à $10^{-6}$ dès $n = 4$. Donc, en calculant les quatre premiers termes de la première série et les deux premiers de la seconde (à $10^{-7}$ près) on obtient

$$\pi \approx 3{,}141592,$$

les cinq premières décimales étant exactes.

Calcul des logarithmes

La série de Taylor de la fonction $y = \ln(1 + x)$ peut être obtenue par intégration de l'identité

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots \quad (|x| < 1),$$

$$\ln(1 + x) = \int_0^x \frac{dx}{1+x} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \tag{4}$$

Pour $x = 1$, cette série converge vers $\ln 2$.

En effet, en intégrant l'identité

$$\frac{1}{1+x} = 1 - x + x^2 - \ldots + (-1)^n x^n + (-1)^{n+1} \frac{x^{n+1}}{1+x}$$

entre 0 et 1, on obtient

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \ldots + (-1)^n \frac{1}{n+1} + \beta_n,$$

où

$$|\beta_n| = \left| \int_0^1 \frac{x^{n+1}}{1+x} dx \right| \leqslant \int_0^1 x^{n+1}\, dx = \frac{1}{n+2} \to 0, \quad n \to \infty.$$

La série (4), de même que la série (3), converge lentement pour $x = 1$.

En substituant $-x$ à $x$ dans (4), on obtient

$$\ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}. \tag{5}$$

En retranchant (5) de (4), on trouve

$$\ln \frac{1+x}{1-x} = 2 \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}. \tag{6}$$

Cette relation est à la base des calculs des logarithmes des entiers naturels. En posant $x = 1/3$ par exemple, on obtient

$$\ln 2 = 2 \sum_{k=0}^{\infty} \frac{1}{(2k+1)\, 3^{2k+1}}, \tag{7}$$

où la série du second membre converge plus vite qu'une progression géométrique.

Pour calculer ln 2 à $10^{-5}$ près il suffit de prendre cinq termes de la série (7) :

$$\ln 2 \approx 0{,}693146$$

(chaque terme de la série est calculé à $10^{-6}$ près).

D'une façon générale, en posant $x = \frac{1}{2m+1}$, où $m$ est un entier naturel, on obtient

$$\frac{1+x}{1-x} = \frac{m+1}{m},$$

$$\ln(m+1) = \ln m + 2 \sum_{k=0}^{\infty} \frac{1}{(2k+1)\,(2m+1)^{2k+1}}. \tag{8}$$

En posant successivement $m = 2, 3, \ldots$, on trouve ln 3, ln 4, ... La série du second membre de (8) converge très vite.

Calcul des racines

Nous avons déjà obtenu la série de Taylor de la fonction $f(x) = (1+x)^\alpha$ au § 16 du chap. 4:

$$(1+x)^\alpha = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!} x^k. \tag{9}$$

La série (9) s'appelle *série binomiale*. On sait que la série (9) ne converge pas toujours pour $x = \pm 1$ et que, si elle converge, elle le fait lentement. Donc, si par exemple l'on a à calculer $\sqrt{2}$, on n'a aucun intérêt à se servir de la formule (9) pour $x = 1$, $\alpha = 1/2$. Il vaut mieux procéder comme suit. On transforme le radicande de telle sorte qu'il diffère peu de l'unité:

$$\sqrt{2} = \sqrt{\frac{2\cdot 25\cdot 49}{25\cdot 49}} = \frac{7}{5}\sqrt{\frac{50}{49}} = \frac{7}{5}\left(1+\frac{1}{49}\right)^{1/2},$$

ou

$$\sqrt{2} = \frac{7}{5}\,\frac{1}{\sqrt{49/50}} = \frac{7}{5}\left(1-\frac{1}{50}\right)^{-1/2} = \frac{7}{5}\left(1-\frac{2}{100}\right)^{-1/2}. \tag{10}$$

Pour trouver les nombres 25 et 49 on agit comme suit: on note les carrés des entiers naturels

$$1,\ 4,\ 9,\ 16,\ 25,\ 36,\ 49,\ 64,\ \ldots, \tag{11}$$

puis, les nombres obtenus en multipliant les carrés (11) par le radicande, c'est-à-dire ici par 2:

$$2,\ 8,\ 18,\ 32,\ 50,\ 72,\ 98,\ 128,\ \ldots \tag{12}$$

Ensuite on prend deux nombres, l'un dans (11) et l'autre dans (12), tels que leur rapport soit proche de l'unité. Il s'agit pour ce qui nous concerne des nombres 49 et 50.

Si l'on prolonge les suites (11) et (12), on peut trouver d'autres nombres dont le quotient est proche de l'unité: par exemple 289 et 288, etc. On a

$$\sqrt{2} = \sqrt{\frac{2\cdot 144\cdot 289}{144\cdot 289}} = \frac{17}{12}\sqrt{\frac{288}{289}} = \frac{17}{12}\left(1+\frac{1}{288}\right)^{-1/2}. \tag{13}$$

On peut se servir maintenant de la série (9). En vertu de (13), pour $x = 1/288$, on obtient

$$\sqrt{2} = \frac{17}{12}\sum_{k=0}^{\infty} \frac{-\frac{1}{2}\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\ldots\left(-\frac{1}{2}-k+1\right)}{k!}\,\frac{1}{288^k}. \tag{14}$$

La série du second membre de (14) converge très vite. De plus, elle est alternée, c'est-à-dire que le reste est inférieur au module du premier membre de ce reste.

Développons la série (14):

$$\sqrt{2}=\frac{17}{12}\left\{1-\frac{1}{2\cdot 288}+\frac{1\cdot 3}{2^2\cdot 2!\,288^2}-\frac{1\cdot 3\cdot 5}{2^3\cdot 3!\,288^3}+\ldots\right\}. \qquad (15)$$

Le troisième terme de la série (15) est inférieur à $8\cdot 10^{-6}<10^{-5}$, donc

$$\sqrt{2}\approx\frac{17}{12}\left(1-\frac{1}{576}\right)=1,414207\ \ldots$$

à $10^{-4}$ près.

Signalons que le calcul de $\sqrt{2}$ à l'aide de (10) est très commode, car au dénominateur on obtient aussitôt des puissances de 10. Si l'on se limite aux trois premiers termes de cette série, on trouve que $\sqrt{2}\approx 1,41421$.

Exemple. Calculer $\sqrt[3]{5}$ à $10^{-2}$ près.

Notons les cubes des entiers naturels

$$1,\ 8,\ 27,\ 64,\ 125,\ 216,\ \ldots$$

et la suite de ces nombres multipliés par 5:

$$5,\ 40,\ 135,\ 320,\ 625,\ 1080,\ \ldots$$

D'où

$$\sqrt[3]{5}=\sqrt[3]{\frac{5\cdot 27\cdot 125}{27\cdot 125}}=\frac{5}{3}\left(1+\frac{10}{125}\right)^{1/3}=\frac{5}{3}\left(1+\frac{8}{100}\right)^{1/3}=$$
$$=\frac{5}{3}\left\{1+\frac{8}{3\cdot 10^2}-\frac{2\cdot 8^2}{3^2\cdot 2!\,10^4}+\frac{2\cdot 5\cdot 8^3}{3^3\cdot 3!\,10^6}-\ldots\right\}.$$

Le troisième terme de la série

$$\frac{5\cdot 8^2}{3^3\cdot 10^4}<0,01,$$

donc

$$\sqrt[3]{5}\approx\frac{5}{3}\left(1+\frac{8}{300}\right)=1,71\ \ldots$$

à $10^{-2}$ près.

## § 15. Séries multiples

On appelle *série double* une expression de la forme

$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}a_{kl}, \qquad (1)$$

où $a_{kl}$ sont des nombres (réels ou complexes) dépendant des indices $k$ et $l = 0, 1, 2, \ldots$ Les nombres $a_{kl}$ s'appellent *termes de la série*, les nombres

$$S_{mn} = \sum_{k=0}^{m} \sum_{l=0}^{n} a_{kl} \quad (m, n = 0, 1, 2, \ldots) \tag{2}$$

*sommes partielles de la série* (1).

Par définition, la série (1) converge vers un nombre $S$, appelé *somme de la série* (1), si existe

$$\lim_{m, n \to \infty} S_{mn} = S, \tag{3}$$

c'est-à-dire si pour tout $\varepsilon > 0$ on peut exhiber un $N$ tel que

$$| S_{mn} - S | < \varepsilon$$

pour tous les $m$, $n > N$. On écrit alors

$$S = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_{kl}.$$

Arrêtons-nous sur le cas où les termes de la série (1) sont positifs ($a_{kl} \geqslant 0$). Posons

$$\Lambda = \sup_{m, n} S_{mn}. \tag{4}$$

Si $\Lambda < \infty$, alors pour tout $\varepsilon > 0$ il existe un couple $(m_0, n_0)$ tel que $\Lambda - \varepsilon < S_{m_0 n_0} \leqslant \Lambda$ et, puisque $a_{kl}$ sont positifs,

$$S_{m_0 n_0} \leqslant S_{mn}, \quad m, \ n > N = \max(m_0, n_0).$$

Donc, $\Lambda - \varepsilon < S_{mn} < \Lambda + \varepsilon$, $m$, $n > N$, et existe

$$\lim_{m, n \to \infty} S_{mn} = S = \Lambda.$$

Si $\Lambda = \infty$, alors de toute évidence

$$\lim_{m, n \to \infty} S_{mn} = S = \infty.$$

On écrit dans ce cas

$$\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_{kl} = \infty.$$

La série (1) est dite *uniformément convergente* si converge la série

$$\sum_{k=0}^{\infty} \sum_{l=0}^{\infty} |a_{kl}|.$$

Comme pour les séries ordinaires, on démontre qu'*une série double absolument convergente est convergente*. La démonstration repose sur

le critère de Cauchy: *pour que la série* (1) *converge il faut et il suffit que pour tout* $\varepsilon > 0$ *il existe un nombre* $N$ *tel que*

$$|S_{mn} - S_{pq}| < \varepsilon$$

*quels que soient* $m, n, p, q > N$.

La démonstration est la même que pour les séries simples.

En plus de la série (1), on peut considérer l'expression

$$\sum_{k=0}^{\infty}\left(\sum_{l=0}^{\infty} a_{kl}\right)$$

à laquelle on associe naturellement le nombre $A$ (s'il existe) obtenu comme suit: si la série entre parenthèses converge pour chaque $k = 0, 1, \ldots$ vers la somme $A_k$ et si la série $\sum_0^{\infty} A_k$ converge vers un nombre $A$, alors on pose

$$A = \sum_{k=0}^{\infty} A_k = \sum_{k=0}^{\infty}\left(\sum_{l=0}^{\infty} a_{kl}\right). \tag{5}$$

THÉORÈME 1. *Si la série* (1) *converge absolument, alors*

$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_{kl} = \sum_{k=0}^{\infty}\left(\sum_{l=0}^{\infty} a_{kl}\right). \tag{6}$$

DÉMONSTRATION. Admettons tout d'abord que $a_{kl} \geqslant 0$. Supposons que le premier membre de (6) (qui a un sens!) est égal à $S$. Pour tous $s$ et $n \geqslant 0$, avec $s \leqslant m$, on a

$$\sum_{l=0}^{n} a_{sl} \leqslant \sum_{k=0}^{m}\sum_{l=0}^{n} a_{kl} \leqslant S, \tag{7}$$

d'où la convergence des séries $\sum_{l=0}^{\infty} a_{sl}$ $(s = 0, 1, \ldots)$. Donc, si dans la deuxième inégalité on fixe $m$ et on passe à la limite pour $n \to \infty$, on obtient

$$\sum_{k=0}^{m}\left(\sum_{l=0}^{\infty} a_{kl}\right) \leqslant S$$

pour tout $m$, d'où l'existence du nombre $A$ (cf. (5)) et le fait que $A \leqslant S$.

Par ailleurs, si $A$ est fini, alors pour tous $m$ et $n$

$$S_{mn} = \sum_{k=0}^{m}\sum_{l=0}^{n} a_{kl} \leqslant \sum_{k=0}^{m}\left(\sum_{l=0}^{\infty} a_{kl}\right) \leqslant A,$$

donc

$$S = \sup_{m,n} S_{mn} \leqslant A.$$

L'égalité (6) est prouvée pour $a_{kl} \geqslant 0$.

Supposons maintenant que $a_{kl}$ sont des nombres quelconques. Posons

$$a_{kl}^{+} = \begin{cases} a_{kl} & (a_{kl} \geqslant 0), \\ 0 & (a_{kl} < 0), \end{cases} \quad a_{kl}^{-} = \begin{cases} -a_{kl} & (a_{kl} \leqslant 0), \\ 0 & (a_{kl} > 0). \end{cases}$$

Alors

$$a_{kl} = a_{kl}^{+} - a_{kl}^{-}, \quad a_{kl}^{+} + a_{kl}^{-} = |a_{kl}|.$$

Donc, la convergence de la série $\sum\sum |a_{kl}|$ entraîne celle des séries $\sum\sum a_{kl}^{+}$, $\sum\sum a_{kl}^{-}$ à termes positifs. Par suite,

$$\sum\sum a_{kl} = \sum\sum a_{kl}^{+} - \sum\sum a_{kl}^{-} =$$
$$= \sum_k \left(\sum_l a_{kl}^{+}\right) - \sum_k \left(\sum_l a_{kl}^{-}\right) = \sum_k \left(\sum_l a_{kl}\right).$$

Enfin, si $a_{kl} = \alpha_{kl} + i\beta_{kl}$ sont des nombres complexes et la série $\sum\sum |a_{kl}|$ converge, alors il en est de même des séries $\sum\sum |\alpha_{kl}|$, $\sum\sum |\beta_{kl}|$, où $\alpha_{kl}$ et $\beta_{kl}$ sont des réels, donc

$$\sum\sum a_{kl} = \sum\sum \alpha_{kl} + i\sum\sum \beta_{kl} =$$
$$= \sum_k \left(\sum_l \alpha_{kl}\right) + i\sum_k \left(\sum_l \beta_{kl}\right) = \sum_k \left(\sum_l a_{kl}\right).$$

Ce qui achève la démonstration du théorème.

EXEMPLE. Pour quelles valeurs de $\alpha > 0$ converge la série double

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{1}{(k+l)^{\alpha}}?$$

SOLUTION. En vertu du théorème 1, nous devons étudier la convergence des séries simples

$$\sum_{l=1}^{\infty} \frac{1}{(k+l)^{\alpha}} = A_k \quad \text{et} \quad \sum_{k=1}^{\infty} A_k.$$

On sait (cf. § 2, théorème 2) que la convergence de la première série équivaut à celle de l'intégrale impropre $\int_0^{\infty} (k+y)^{-\alpha}\, dy$ qui converge pour $\alpha > 1$:

$$A_k = \sum_{l=1}^{\infty} (k+l)^{-\alpha} \leqslant \int_0^{\infty} (k+y)^{-\alpha}\, dy = \frac{(k+y)^{1-\alpha}}{1-\alpha}\Big|_0^{\infty} = \frac{k^{1-\alpha}}{\alpha-1} \quad (\alpha > 1).$$

D'autre part,

$$\sum_{k=1}^{\infty} A_k \leqslant \frac{1}{\alpha-1} \sum_{k=1}^{\infty} k^{1-\alpha} = \frac{1}{\alpha-1} + \sum_{k=2}^{\infty} k^{1-\alpha} \leqslant \frac{1}{\alpha-1} + \int_1^{\infty} x^{1-\alpha}\,dx.$$

La dernière intégrale converge pour $\alpha - 1 > 1$, c'est-à-dire pour $\alpha > 2$. Donc, la série double envisagée converge pour $\alpha > 2$.

Considérons maintenant le problème suivant. Soit donnée une série double (1) convergeant absolument. Sa somme $S$, de même que la somme $S'$ de la série des modules de ses termes peuvent s'écrire sous la forme de limites de suites:

$$S = \lim_{n\to\infty} \sum_0^n \sum_0^n a_{kl} = \lim_{n\to\infty} S_{nn},$$

$$S' = \lim_{n\to\infty} \sum_0^n \sum_0^n |a_{kl}| = \lim_{n\to\infty} S'_{nn}$$

dépendant d'un seul indice $n$. Aux suites $\{S_{nn}\}$ et $\{S'_{nn}\}$ correspondent les séries convergentes

$$S = a_{00} + (a_{10} + a_{11} + a_{01}) + (a_{20} + a_{21} + a_{22} + a_{12} + a_{02}) + (a_{30} + \ldots) + \ldots, \qquad (8)$$

$$S' = |a_{00}| + (|a_{10}| + |a_{11}| + |a_{01}|) + (|a_{20}| + |a_{21}| + |a_{22}| + |a_{12}| + |a_{02}|) + (|a_{30}| + \ldots) + \ldots \qquad (9)$$

dont les termes sont égaux aux sommes entre parenthèses. Or, les parenthèses de la deuxième série sont positives, donc on ne modifie pas la convergence en les supprimant:

$$|a_{00}| + |a_{10}| + |a_{11}| + |a_{01}| + |a_{20}| + \ldots \qquad (10)$$

Donc, la série

$$a_{00} + a_{10} + a_{11} + a_{01} + a_{20} + \ldots, \qquad (11)$$

obtenue à partir de (8) par suppression de toutes les parenthèses, converge absolument. Par suite, elle converge visiblement vers $S$.

Nous avons montré que si la série double (1) converge vers un nombre $S$ absolument, alors la série simple (11) converge aussi absolument vers $S$. Or, on peut changer l'ordre des termes d'une série absolument convergente sans modifier la convergence et la somme.

Nous venons donc de prouver le

THÉORÈME 2. *La série $v_0 + v_1 + v_2 + \ldots$, obtenue en numérotant à l'aide d'un seul indice les termes de la série* (1) *absolument convergente vers un nombre $S$, converge de même absolument vers $S$.*

Prouvons le théorème suivant en conclusion.

THÉORÈME 3. *Soient données deux séries absolument convergentes* $\sum_0^\infty u_k$, $\sum_0^\infty v_l$. *Supposons que les produits* $u_k v_l$ ($k$, $l = 0, 1, 2, \ldots$) *sont numérotés par un seul indice de la manière suivante*: $w_0, w_1, w_2, \ldots$ *On a alors*

$$\sum_0^\infty u_k \times \sum_0^\infty v_l = \sum_0^\infty w_k,$$

*où la série du second membre converge absolument.*

DÉMONSTRATION. En effet, supposons que $\sum_0^\infty |u_k| = M$, $\sum_0^\infty |v_l| = N$. La série double $\sum_{k=0}^\infty \sum_{l=0}^\infty u_k v_l$ converge absolument, car pour tous $m$ et $n$

$$\sum_0^m \sum_0^n |u_k v_l| = \sum_0^m |u_k| \times \sum_0^n |v_l| \leqslant MN.$$

Donc

$$\sum_{k=0}^\infty u_k \times \sum_{l=0}^\infty v_l = \lim_{n\to\infty} \sum_{k=0}^n u_k \times \lim_{n\to\infty} \sum_{l=0}^n v_l =$$

$$= \lim_{n\to\infty} \sum_{k=0}^n \sum_{l=0}^n u_k v_l = \sum_{k=0}^\infty \sum_{l=0}^\infty u_k v_l = \sum_{j=0}^\infty w_j,$$

où la dernière égalité a lieu grâce au théorème précédent.

En conclusion de ce paragraphe signalons qu'on peut envisager des séries triples, quadruples et plus généralement $n$-tuples:

$$\sum_{k=0}^\infty \sum_{l=0}^\infty \sum_{m=0}^\infty a_{klm}, \ \ldots, \ \sum_{k_1=0}^\infty \ldots \sum_{k_n=0}^\infty a_{k_1 k_2 \ldots k_n}.$$

Ces séries sont justiciables de théorèmes analogues aux théorèmes 1, 2 et 3.

EXERCICES.

1. Pour quelles valeurs de $\alpha > 0$ converge la série triple

$$\sum_{k=1}^\infty \sum_{l=1}^\infty \sum_{m=1}^\infty \frac{1}{(k+l+m)^\alpha}.$$

Réponse: pour $\alpha > 3$.

2. Etudier la convergence de la série $n$-tuple

$$\sum_{k_1=1}^{\infty} \cdots \sum_{k_n=1}^{\infty} \frac{1}{(k_1+\ldots+k_n)^{\alpha}}.$$

Réponse: $\alpha > n$.

3. Pour quelles valeurs de $\alpha > 0$ converge la série double

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{1}{(k^2+l^2)^{\alpha}}?$$

Réponse: $\alpha > 1$.

4. Pour quelles valeurs de $\alpha$, $\beta_1$ et de $\beta_2$ converge la série double

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{1}{(k^{\beta_1}+l^{\beta_2})^{\alpha}}?$$

Réponse: $\alpha > \frac{1}{\beta_1} + \frac{1}{\beta_2}$.

5. Pour quelles valeurs de $\alpha$, $\beta_1$, $\beta_2$, $\ldots$, $\beta_n > 0$ converge la série $n$-tuple

$$\sum_{k_1=1}^{\infty} \cdots \sum_{k_n=1}^{\infty} \frac{1}{(k_1^{\beta_1}+\ldots+k_n^{\beta_n})^{\alpha}}?$$

Réponse: $\alpha > \sum_{j=1}^{n} \frac{1}{\beta_j}$.

## § 16. Sommation des séries et suites par la méthode des moyennes arithmétiques

Soit donnée la série numérique

$$u_0 + u_1 + u_2 + \ldots \tag{1}$$

Posons

$$S_n = u_0 + u_1 + \ldots + u_n, \tag{2}$$

$$\sigma_n = \frac{S_0+S_1+\ldots+S_n}{n+1} \quad (n=0,\ 1,\ 2\ \ldots). \tag{3}$$

Par définition, *la série* (1) (*ou la suite* $\{S_n\}$) *est sommable par la méthode des moyennes arithmétiques au nombre* $\sigma$ *si existe la limite*

$$\lim_{n\to\infty} \sigma_n = \sigma. \tag{4}$$

THÉORÈME. *Si la série (1) converge vers un nombre $S$, alors elle est sommable par la méthode des moyennes arithmétiques et cette somme est égale à $S$.*

DÉMONSTRATION. Supposons que la série (1) converge ; il existe alors un nombre $M > 0$ tel que

$$|S_j| \leqslant M \quad (j = 0, 1, 2, \ldots), \tag{5}$$

et un entier naturel $n$ assez grand, que nous supposerons fixe (un nombre $k$ et, dans la suite, un nombre $p$, variables) tel que

$$|S_{n+k} - S| < \varepsilon \quad (k = 1, 2, \ldots). \tag{6}$$

On a par ailleurs

$$S - \sigma_{n+p} = \left(S - \frac{1}{p}\sum_{k=1}^{p} S_{n+k}\right) + \\ + \left(\frac{1}{p} - \frac{1}{n+p+1}\right)\sum_{k=1}^{p} S_{n+k} - \frac{1}{n+p+1}\sum_{k=0}^{n} S_k = \\ = \frac{1}{p}\sum_{k=1}^{p}(S - S_{n+k}) + \left(\frac{1}{p} - \frac{1}{n+p+1}\right)\sum_{k=1}^{p} S_{n+k} - \frac{1}{n+p+1}\sum_{k=0}^{n} S_k,$$

d'où, puisque

$$\frac{1}{p} - \frac{1}{n+p+1} = \frac{n+1}{p(n+p+1)},$$

l'on déduit

$$|S - \sigma_{n+p}| < \varepsilon + \frac{n+1}{n+p+1}M + \\ + \frac{n+1}{n+p+1}M < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \quad (p > p_0),$$

si $p$ est assez grand. Par suite, $\sigma_{n+p} \to S$ $(p \to \infty)$ ou, ce qui revient au même, $\sigma_j \to S$ $(j \to \infty)$, ce qui prouve le théorème.

EXEMPLE 1. Considérons la série $\sum_{n=0}^{\infty} 2^{-n} = 2$. Ici

$$S_0 = 1,\ S_1 = 1 + \frac{1}{2},\ \ldots,\ S_n = 1 + \frac{1}{2} + \ldots + \frac{1}{2^n} = 2 - \frac{1}{2^n}.$$

D'où

$$\sigma_n = \frac{S_0 + S_1 + \ldots + S_n}{n+1} = \frac{\sum_{k=0}^{n}\left(2 - \frac{1}{2^k}\right)}{n+1} = \\ = \frac{2(n+1) - 1 + \frac{1}{2^n}}{n+1} = 2 - \frac{1}{n+1} + \frac{1}{2^n(n+1)} \to 2 \quad (n \to \infty).$$

Exemple 2. La série $1-1+1-\ldots$ diverge, mais par la méthode des moyennes arithmétiques sa somme est égale à $\frac{1}{2}$.

En effet, $S_0=1,\ S_1=0,\ \ldots,\ S_{2n}=1,\ S_{2n+1}=0,\ \ldots$ Donc

$$\sigma_{2n}=\frac{S_0+\ldots+S_{2n}}{2n+1}=\frac{1+0+\ldots+0+1}{2n+1}=\frac{n+1}{2n+1}\rightarrow\frac{1}{2},\quad n\rightarrow\infty;$$

$$\sigma_{2n+1}=\frac{S_0+\ldots+S_{2n}+S_{2n+1}}{2n+2}=\frac{n+1}{2(n+1)}=\frac{1}{2}\rightarrow\frac{1}{2},\quad n\rightarrow\infty.$$

CHAPITRE 10

# ÉLÉMENTS D'ALGÈBRE LINÉAIRE ET DE GÉOMÉTRIE ANALYTIQUE

## § 1. Déterminants du second ordre

Soient donnés des nombres réels ou complexes $a_1$, $a_2$, $b_1$, $b_2$. Le nombre $a_1b_2 - a_2b_1$ s'appelle *déterminant du second ordre* et se note :

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1. \tag{1}$$

Les nombres $a_1$, $a_2$, $b_1$, $b_2$ sont les *éléments du déterminant*. Les nombres $a_1$, $a_2$ forment la *première ligne*, les nombres $b_1$, $b_2$ la *deuxième ligne*, les nombres $a_1$, $b_1$, la *première colonne*, les nombres $a_2$, $b_2$, la *deuxième colonne*.

Les *propriétés* suivantes *du déterminant* sont de vérification immédiate.

*Le déterminant*

a) *ne change pas de valeur si l'on échange les lignes et les colonnes* :

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix};$$

b) *change de signe si on permute deux lignes ou deux colonnes* :

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = -\begin{vmatrix} b_1 & b_2 \\ a_1 & a_2 \end{vmatrix},$$

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = -\begin{vmatrix} a_2 & a_1 \\ b_2 & b_1 \end{vmatrix};$$

c) *est multiplié par k si les éléments d'une ligne ou d'une colonne sont multipliés par k.* Ainsi

$$\begin{vmatrix} ka_1 & b_1 \\ ka_2 & b_2 \end{vmatrix} = k\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix},$$

24—0622

autrement dit *si les éléments d'une ligne ou d'une colonne sont proportionnels, on peut sortir le coefficient de proportionnalité du signe du déterminant*;

d) *est nul si les éléments d'une ligne ou d'une colonne sont tous nuls.* Par exemple

$$\begin{vmatrix} 0 & 0 \\ b_1 & b_2 \end{vmatrix} = 0b_2 - 0b_1 = 0;$$

e) *est nul s'il a deux lignes ou deux colonnes identiques:*

$$\begin{vmatrix} a_1 & a_1 \\ a_2 & a_2 \end{vmatrix} = a_1a_2 - a_1a_2 = 0.$$

Dans la suite on introduit les déterminants du troisième et du $n$-ième ordre. Ces déterminants sont justiciables des propriétés a), b), c), d), e).

## § 2. Déterminants du troisième et du $n$-ième ordre

On appelle *déterminant du troisième ordre* le nombre

$$\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \quad (1)$$

écrit sous la forme

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \quad (2)$$

Le nombre $a_{kl}$ s'appelle *élément* du déterminant. Cet élément se trouve à l'intersection de la $k$-ième ligne et de la $l$-ième colonne. La diagonale $a_{11}$, $a_{22}$, $a_{33}$ s'appelle *diagonale principale*, la diagonale $a_{13}$, $a_{22}$, $a_{31}$, *diagonale non principale.*

L'expression (1) est d'une structure assez simple. C'est le nombre calculé avec les éléments $a_{kl}$ d'après la règle suivante dite de Sarrus: formons le tableau de Sarrus en ajoutant les deux premières colonnes à droite du déterminant (2) (voir fig. 110). On voit que le nombre $\Delta$ est égal à la somme des produits des éléments reliés par une droite, ces produits étant pris avec le signe + ou — selon que cette droite est parallèle à la diagonale principale ou à la diagonale non principale.

Chacun de ces produits signés s'appelle *terme du déterminant* (2). Chaque produit est composé d'un élément de chaque ligne et de chaque colonne. Dans chaque terme de la somme (1) ces éléments sont classés par ordre de croissance du premier indice, c'est-à-dire

l'indice des lignes. Les indices des colonnes sont disposés comme suit :

$$\begin{cases} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{cases} \tag{3}$$

$$\begin{cases} 3 & 2 & 1 \\ 1 & 3 & 2 \\ 2 & 1 & 3 \end{cases} \tag{4}$$

On reconnaît toutes les permutations des nombres 1, 2, 3. La permutation

$$1\ 2\ 3 \tag{5}$$

des nombres 1, 2, 3 sera dite *permutation de base.*

On appelle *transposition de deux éléments d'une permutation* l'opération qui consiste à échanger ces deux éléments. La transposition transforme une permutation en une permutation. Si dans la

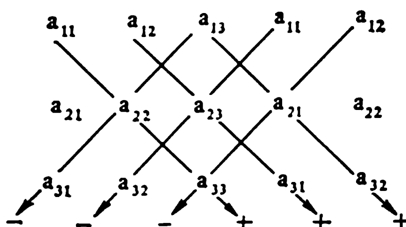



Fig. 110

dernière permutation on effectue une transposition, on obtient une troisième permutation (qui peut très bien coïncider avec la première). Par exemple, la permutation

$$3\ 2\ 1 \tag{6}$$

se déduit de (5) par transposition du premier et du troisième élément, la permutation

$$2\ 3\ 1 \tag{7}$$

se déduit de (6) par transposition du premier et du deuxième élément.

Il est important de noter que si une permutation se déduit à partir de la permutation de base par $N$ transpositions et si la même permutation s'obtient à partir de la permutation de base d'une autre manière par $N_1$ transpositions, les nombres $N$ et $N_1$ sont simultanément pairs ou impairs. Une permutation des nombres 1, 2, 3

est dite *paire* (resp. *impaire*) si elle se déduit de la permutation de base par un nombre pair (resp. impair) de transpositions.

Soit donnée une permutation $j = (j_1, j_2, j_3)$, où $j_1, j_2, j_3$ sont les nombres 1, 2, 3 pris dans un ordre quelconque. Désignons par $t(j)$ le *nombre de transpositions* nécessaires pour déduire cette permutation de la permutation de base. La permutation $j$ est paire (resp. impaire) si $t(j)$ est pair (resp. impair).

Les permutations (3) sont paires, les permutations (4), impaires.

A la lumière de ce qui vient d'être dit on peut donner une définition équivalente du déterminant d'ordre trois.

On appelle *déterminant d'ordre trois* (2) le nombre Δ égal à la somme

$$\Delta = \sum_j (-1)^{t(j)} a_{1j_1} a_{2j_2} a_{3j_3} \tag{8}$$

des produits de la forme $(-1)^{t(j)} a_{1j_1} a_{2j_2} a_{3j_3}$, où $j = (j_1, j_2, j_3)$ représentent toutes les permutations des nombres 1, 2, 3.

Cette définition se généralise à un déterminant d'ordre $n$ ($n = 1, 2, 3, \ldots$).

On appelle *déterminant d'ordre n* le nombre

$$\Delta = |a_{ik}| = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}, \tag{9}$$

égal à la somme

$$\Delta = \sum_j (-1)^{t(j)} a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

étendue à toutes les permutations $j = (j_1, \ldots, j_n)$. Le nombre $t(j)$ s'appelle *nombre de transpositions* pour passer de la permutation de base $1, 2, \ldots, n$ à la permutation $j = (j_1, \ldots, j_n)$. Le produit $(-1)^{t(j)} a_{1j_1} \ldots a_{nj_n}$ s'appelle *terme du déterminant*.

*Les déterminants d'ordre n possèdent les propriétés* a), b), c), d), e) *énumérées dans le paragraphe précédent.*

DÉMONSTRATION. a) Si l'on échange les lignes et les colonnes, les indices des lignes viendront en deuxième position. Par exemple, pour le déterminant d'ordre trois (2) on aura

$$\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - \\ - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} = \Delta.$$

Le terme général du nouveau déterminant d'ordre $n$ s'écrit

$$(-1)^{t(j)} a_{j_1 1} a_{j_2 2} \ldots a_{j_n n}.$$

Ordonnons les facteurs du produit $a_{j_1 1} a_{j_2 2} \ldots a_{j_n n}$ par rapport au premier indice, c'est-à-dire qu'on passe de la permutation $j = (j_1, \ldots, j_n)$ à la permutation de base 1, 2, . . ., $n$. Pour cela il faut $t(j)$ transpositions. La permutation de base des deuxièmes indices se transforme alors en une permutation $i = (i_1, \ldots, i_n)$, et le nombre $t(i)$ sera de même parité que le nombre $t(j)$. Donc

$$(-1)^{t(j)} a_{j_1 1} \ldots a_{j_n n} = (-1)^{t(i)} a_{1 i_1} \ldots a_{n i_n}.$$

Il est immédiat de voir qu'à des permutations distinctes $j_1, \ldots, j_n$ correspondent des permutations distinctes $i_1, \ldots, i_n$. Donc

$$\sum_j (-1)^{t(j)} a_{j_1 1} \ldots a_{j_n n} = \sum_i (-1)^{t(i)} a_{1 i_1} \ldots a_{n i_n} = \Delta.$$

b) Echangeons par exemple la première et la troisième ligne du déterminant d'ordre trois (2). Le déterminant $\Delta'$ obtenu sera égal à

$$\Delta' = \begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{vmatrix} = \sum_j (-1)^{t(j)} a_{3j_1} a_{2j_2} a_{1j_3} =$$

$$= \sum_j (-1)^{t(j)} a_{1j_3} a_{2j_2} a_{3j_1} = - \sum_{j'=(j_3, j_2, j_1)} (-1)^{t(j')} a_{1j_3} a_{2j_2} a_{3j_1} = -\Delta,$$

puisque la permutation $j = (j_1, j_2, j_3)$ se déduit de la permutation $j' = (j_3, j_2, j_1)$ par une seule transposition.

c) La multiplication d'une ligne (resp. colonne) quelconque par un nombre $k$ revient à multiplier tous les termes du déterminant par $k$, puisque chaque terme contient un élément de la ligne (resp. colonne) mentionnée. Donc, la somme des termes est multipliée par $k$.

d) Un déterminant dont les éléments d'une ligne ou d'une colonne sont nuls est nul, puisque tous ses termes le sont de toute évidence.

e) Un déterminant est nul s'il a deux lignes ou deux colonnes identiques. Ceci résulte de la propriété b) ($\Delta' = -\Delta$, $\Delta' = \Delta$, d'où $\Delta = 0$).

Supprimons la $i$-ième ligne et la $k$-ième colonne du déterminant d'ordre $n$ (9). L'expression restante engendre un déterminant $M_{ik}$ d'ordre $n - 1$ appelé *mineur de l'élément* $a_{ik}$. La quantité

$$A_{ik} = (-1)^{i+k} M_{ik}$$

s'appelle *cofacteur* de $a_{ik}$.

PROPRIÉTÉ f) *La somme des produits des éléments $a_{ik}$ d'une ligne (resp. colonne) par leurs cofacteurs est égale à* $\Delta$:

$$\Delta = \sum_{k=1}^{n} a_{ik}A_{ik} \quad (i = 1, \ldots, n), \tag{10}$$

$$\Delta = \sum_{i=1}^{n} a_{ik}A_{ik} \quad (k = 1, \ldots, n). \tag{10'}$$

Prouvons cette propriété pour la troisième ligne d'un déterminant d'ordre trois. On a

$$a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33} =$$

$$= a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} =$$

$$= a_{31}(a_{12}a_{23} - a_{13}a_{22}) + a_{32}(a_{13}a_{21} - a_{11}a_{23}) +$$

$$+ a_{33}(a_{11}a_{22} - a_{12}a_{21}) = \Delta.$$

La somme (10) s'appelle *développement du déterminant suivant les éléments de la i-ième ligne*, la somme (10'), *développement du déterminant suivant les éléments de la k-ième colonne.*

EXEMPLE 1. Si dans le déterminant $\Delta$ (voir (9)) on a $a_{21} = a_{31} = \ldots = a_{n1} = 0$, alors $\Delta = a_{11}A_{11}$, c'est-à-dire que le calcul de ce déterminant se ramène à celui d'un de ses cofacteurs qui est un déterminant d'ordre $n - 1$.

EXEMPLE 2. Si tous les éléments sous-diagonaux (resp. subdiagonaux) de $\Delta$ sont nuls ($a_{kl} = 0$ pour $k > l$ (resp. $k < l$)), alors $\Delta = a_{11}a_{22} \ldots a_{nn}$. Ceci résulte de l'exemple précédent.

PROPRIÉTÉ g) *La somme des produits des éléments $a_{ik}$ d'une ligne (resp. d'une colonne) d'un déterminant par les cofacteurs respectifs des éléments d'une autre ligne (resp. colonne) est nulle*:

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \sum_{k=1}^{n} a_{ki}A_{kj} = 0 \tag{11}$$

$(i \neq j;\ i, j = 1, \ldots, n)$.

En effet, considérons la première somme. On voit qu'elle ne dépend pas des éléments de la $j$-ième ligne. L'échange de la $j$-ième ligne et de la $i$-ième ligne ne modifie pas cette somme. Mais on peut désormais la considérer comme le développement du nouveau déterminant suivant les éléments de la $j$-ième ligne. Or, ce développement est nul, car le déterminant a deux lignes identiques, la $i$-ième et la $j$-ième.

PROPRIÉTÉ h) *Soient donnés deux déterminants $\Delta_1$ et $\Delta_2$ d'ordre n dont toutes les lignes (resp. colonnes) sont identiques à l'exception d'une, la i-ième pour fixer les idées. La somme de ces déterminants est égale à un déterminant d'ordre n dont toutes les lignes sont identiques à celles de $\Delta_1$ et de $\Delta_2$ à l'exception de la i-ième ligne (resp. colonne) qui est formée de la somme des éléments respectifs des i-ièmes lignes (resp. colonnes) des déterminants $\Delta_1$ et $\Delta_2$.* Ainsi

$$\Delta_1+\Delta_2=\begin{vmatrix} a_{11} & \dots & a_{1,n-1} & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{n,n-1} & a_{nn} \end{vmatrix}+\begin{vmatrix} a_{11} & \dots & a_{1,n-1} & b_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{n,n-1} & b_{nn} \end{vmatrix}=$$

$$=\begin{vmatrix} a_{11} & \dots & a_{1,n-1} & (a_{1n}+b_{1n}) \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{n,n-1} & (a_{nn}+b_{nn}) \end{vmatrix}=\Delta.$$

En effet, un développement de ces déterminants suivant les éléments de la $n$-ième colonne donne

$$\Delta_1+\Delta_2=\sum_{k=1}^{n} a_{kn}A_{kn}+\sum_{k=1}^{n} b_{kn}A_{kn}=\sum_{k=1}^{n}(a_{kn}+b_{kn})A_{kn}=\Delta.$$

PROPRIÉTÉ i) *Un déterminant ne change pas de valeur si aux éléments d'une ligne (resp. colonne) on ajoute les éléments respectifs d'une autre ligne (resp. colonne) multipliés par un nombre k.* Par exemple

$$\begin{vmatrix} a_{11} & \dots & a_{1,n-1} & a_{1n}+ka_{11} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{n,n-1} & a_{nn}+ka_{n1} \end{vmatrix}=$$

$$=\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} \end{vmatrix}+k\begin{vmatrix} a_{11} & \dots & a_{1,n-1} & a_{11} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{n,n-1} & a_{n1} \end{vmatrix}=$$

$$=|a_{kl}|+k\cdot 0=|a_{kl}|,$$

en vertu des propriétés h), c), e).

L'utilisation adéquate de la propriété i) ramène le calcul d'un déterminant à celui d'un autre d'ordre moindre.

EXEMPLE 3.

$$\begin{vmatrix} 1 & 4 & 5 \\ 2 & 3 & 1 \\ 8 & 1 & 1 \end{vmatrix}=\begin{vmatrix} 1 & 4 & 5 \\ 0 & -5 & -9 \\ 8 & 1 & 1 \end{vmatrix}=\begin{vmatrix} 1 & 4 & 5 \\ 0 & -5 & -9 \\ 0 & -31 & -39 \end{vmatrix}=$$

$$=\begin{vmatrix} -5 & -9 \\ -31 & -39 \end{vmatrix}=\begin{vmatrix} 5 & 9 \\ 31 & 39 \end{vmatrix}=5\cdot 39-9\cdot 31=-84.$$

EXEMPLE 4.

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 0 \\ -5 & 3 & -1 \\ -7 & 4 & 2 \end{vmatrix} = -\begin{vmatrix} -5 & -1 \\ -7 & 2 \end{vmatrix} = \begin{vmatrix} 5 & 1 \\ -7 & 2 \end{vmatrix} = 17.$$

EXEMPLE 5. Le déterminant

$$\Delta_n = \begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{vmatrix}$$

s'appelle ***déterminant de Vandermonde*** [1]).

Ce déterminant est nul si deux nombres quelconques $a_j$ et $a_k$ sont égaux. Si les $a_j$ sont tous distincts, alors

$$\begin{aligned} \Delta_n = {} & (a_n - a_1)(a_{n-1} - a_1) \dots (a_2 - a_1) \times \\ & \times (a_n - a_2)(a_{n-1} - a_2) \dots (a_3 - a_2) \times \\ & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ & \times (a_n - a_{n-2})(a_{n-1} - a_{n-2}) \times \\ & \times (a_n - a_{n-1}). \end{aligned} \qquad (12)$$

Prouvons cette formule par récurrence. Pour $n = 2$ on a

$$\Delta_2 = \begin{vmatrix} 1 & a_1 \\ 1 & a_2 \end{vmatrix} = a_2 - a_1.$$

Supposons que cette formule est vraie pour $n = k - 1$ et montrons qu'elle l'est pour $n = k$. Utilisons les propriétés i) et c). Multiplions la $(k - 1)$-ième colonne de $\Delta_k$ par $a_1$ et retranchons-la de la $k$-ième, multiplions la $(k - 2)$-ième colonne par $a_1$ et retranchons-la de la $(k - 1)$-ième et ainsi de suite; on obtient

$$\Delta_k = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & a_2 - a_1 & a_2(a_2 - a_1) & \dots & a_2^{k-2}(a_2 - a_1) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & a_k - a_1 & a_k(a_k - a_1) & \dots & a_k^{k-2}(a_k - a_1) \end{vmatrix} =$$

$$= (a_k - a_1)(a_{k-1} - a_1) \dots (a_2 - a_1) \begin{vmatrix} 1 & a_2 & \dots & a_2^{k-2} \\ 1 & a_3 & \dots & a_3^{k-2} \\ \cdot & \cdot & \cdot & \cdot \\ 1 & a_k & \dots & a_k^{k-2} \end{vmatrix}.$$

[1]) Alexandre Théophile Vandermonde, mathématicien français Paris) 1735, id. 1796).

Le dernier déterminant est un déterminant de Vandermonde d'ordre $k - 1$ engendré par les nombres $a_2, \ldots, a_k$. Donc, on a par hypothèse

$$\begin{aligned}\Delta_k = &(a_k - a_1)(a_{k-1} - a_1) \ldots (a_2 - a_1) \times \\ &\times (a_k - a_2)(a_{k-1} - a_2) \ldots (a_3 - a_2) \times \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\times (a_k - a_{k-2})(a_{k-1} - a_{k-2}) \times \\ &\times (a_k - a_{k-1}),\end{aligned}$$

c.q.f.d.

PROPRIÉTÉ j) *Soient*

$$\Delta_1 = |b_{kl}|, \quad \Delta_2 = |a_{kl}|.$$

*Le produit de deux déterminants* $|b_{kl}|$ *et* $|a_{kl}|$ *d'ordre n est un déterminant d'ordre n:*

$$\Delta = \left|\sum_{j=1}^{n} b_{kj}a_{jl}\right|.$$

Donc, l'élément $\gamma_{kl}$ de $\Delta$ est égal au *produit de la k-ième ligne du déterminant* $\Delta_1$ *par la l-ième colonne du déterminant* $\Delta_2$, ou plus exactement à la *somme des produits* des éléments de la $k$-ième ligne de $\Delta_1$ par ceux de la $l$-ième colonne de $\Delta_2$.

Puisque dans les déterminants $\Delta_1$ et $\Delta_2$ on peut échanger les lignes et les colonnes, il est évident que les éléments $\gamma_{kl}$ de $\Delta$ sont égaux soit au produit de la $k$-ième ligne de $\Delta_1$ par la $l$-ième ligne de $\Delta_2$, soit au produit de la $k$-ième colonne de $\Delta_1$ par la $l$-ième colonne de $\Delta_2$, soit enfin au produit de la $k$-ième colonne de $\Delta_1$ par la $l$-ième ligne de $\Delta_2$.

DÉMONSTRATION. Vérifions cette propriété sur l'exemple des déterminants d'ordre deux:

$$\Delta_1 = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta = \begin{vmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{vmatrix},$$

où

$$\gamma_{11} = b_{11}a_{11} + b_{12}a_{21}, \quad \gamma_{12} = b_{11}a_{12} + b_{12}a_{22},$$
$$\gamma_{21} = b_{21}a_{11} + b_{22}a_{21}, \quad \gamma_{22} = b_{21}a_{12} + b_{22}a_{22}.$$

D'après les propriétés h), c), e)

$$\Delta = \begin{vmatrix} b_{11}a_{11} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} & b_{21}a_{12} + b_{22}a_{22} \end{vmatrix} + \begin{vmatrix} b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{vmatrix} =$$

$$=a_{11}a_{12}\begin{vmatrix} b_{11} & b_{11} \\ b_{21} & b_{21} \end{vmatrix}+a_{11}a_{22}\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}+$$

$$+a_{12}a_{21}\begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{vmatrix}+a_{21}a_{22}\begin{vmatrix} b_{12} & b_{12} \\ b_{22} & b_{22} \end{vmatrix}=$$

$$=a_{11}a_{12}\cdot 0+a_{11}a_{22}\Delta_1+a_{12}a_{21}\begin{vmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{vmatrix}+a_{21}a_{22}\cdot 0=$$

$$=a_{11}a_{22}\Delta_1-a_{12}a_{21}\Delta_1=\Delta_1(a_{11}a_{22}-a_{12}a_{21})=\Delta_1\Delta_2.$$

Dans le cas d'un déterminant d'ordre $n$ on peut écrire

$$\Delta=\begin{vmatrix} \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_1 1} & \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_1 2} & \cdots & \sum\limits_{s_1=1}^{n} a_{1s_1}b_{s_1 n} \\ \sum\limits_{s_2=1}^{n} a_{2s_2}b_{s_2 1} & \sum\limits_{s_2=1}^{n} a_{2s_2}b_{s_2 2} & \cdots & \sum\limits_{s_2=1}^{n} a_{2s_2}b_{s_2 n} \\ \cdot & \cdot & \cdot & \cdot \\ \sum\limits_{s_n=1}^{n} a_{ns_n}b_{s_n 1} & \sum\limits_{s_n=1}^{n} a_{ns_n}b_{s_n 2} & \cdots & \sum\limits_{s_n=1}^{n} a_{ns_n}b_{s_n n} \end{vmatrix}=$$

$$=\sum_{s_1=1}^{n}\sum_{s_2=1}^{n}\cdots\sum_{s_n=1}^{n} a_{1s_1}a_{2s_2}\cdots a_{ns_n}\begin{vmatrix} b_{s_1 1} & b_{s_1 2} & \cdots & b_{s_1 n} \\ b_{s_2 1} & b_{s_2 2} & \cdots & b_{s_2 n} \\ \cdot & \cdot & \cdot & \cdot \\ b_{s_n 1} & b_{s_n 2} & \cdots & b_{s_n n} \end{vmatrix}=$$

$$=\sum_{s=(s_1,\, s_2,\, \ldots,\, s_n)} a_{1s_1}a_{2s_2}\cdots a_{ns_n}(-1)^{t(s)}\Delta_1=\Delta_2\Delta_1.$$

Pour calculer les éléments de $\Delta$ on peut prendre n'importe quel indice de sommation $s$ $\left(\gamma_{kl}=\sum\limits_{s=1}^{n} a_{ks}b_{sl}\right)$, mais pour la commodité on se servira de l'indice $s_1$ pour la première ligne de $\Delta$, de l'indice $s_2$ pour la deuxième ligne, etc. La deuxième égalité est réalisée en vertu des propriétés h) et c); la somme multiple $\sum\limits_{s_1=1}^{n}\sum\limits_{s_2=1}^{n}\cdots\sum\limits_{s_n=1}^{n}$ est étendue à toutes les permutations $(s_1, s_2, \ldots, s_n)$, où $1 \leqslant s_j \leqslant n$. Cependant, si dans un système quelconque $(s_1, s_2, \ldots, s_n)$ deux composantes $s_i$ et $s_j$ sont égales $(s_i = s_j,\ i \neq j)$, alors le déterminant $|b_{s_k l}| = 0$. Donc, dans la somme multiple on ne peut conserver que les termes correspondant aux permutations distinctes $(s_1, \ldots, s_n)$ des entiers naturels $(1, \ldots, n)$. Ceci étant, il est évident que

$$|b_{s_k l}| = (-1)^{t(s)}\Delta_1.$$

## § 3. Matrices

Un tableau de nombres (réels ou complexes) $\alpha_{ij}$ de la forme

$$A=\begin{Vmatrix}\alpha_{11} & \dots & \alpha_{1n}\\ \cdot & \cdot\cdot\cdot & \cdot\\ \alpha_{m1} & \dots & \alpha_{mn}\end{Vmatrix}=\begin{pmatrix}\alpha_{11} & \dots & \alpha_{1n}\\ \cdot & \cdot\cdot\cdot & \cdot\\ \alpha_{m1} & \dots & \alpha_{mn}\end{pmatrix}=\|\alpha_{ij}\|=(\alpha_{ij}) \qquad (1)$$

s'appelle *matrice* à $m$ lignes et $n$ colonnes. Les nombres $\alpha_{ij}$ sont les *éléments* de la matrice. La matrice considérée est rectangulaire. Pour $m = n$ c'est une *matrice carrée* d'ordre $n$.

Une matrice $B = \| \beta_{ij} \|$ à $m$ lignes et $n$ colonnes est par définition *égale* à la matrice $A$ si et seulement si les éléments correspondants de ces matrices sont égaux ($\alpha_{ij} = \beta_{ij}$). On écrit alors $A = B$. La matrice $\| \alpha_{ij} \|$ n'est pas un nombre, c'est un tableau. Cependant pour une matrice carrée on peut considérer le nombre $| \alpha_{ij} |$ ou *déterminant engendré par cette matrice*.

Soit $k$ ($\leqslant m, n$) un entier naturel. Supprimons $k$ colonnes et $k$ lignes quelconques du tableau (1). Les éléments $\alpha_{js}$ situés à l'intersection des colonnes et des lignes supprimées forment une matrice carrée qui engendre un déterminant d'ordre $k$. Ce déterminant s'appelle *déterminant d'ordre k engendré par la matrice A*.

On appelle *rang* d'une matrice $A$ le plus grand entier naturel $k$ pour lequel il existe un déterminant non nul d'ordre $k$ engendré par la matrice $A$ (cf. remarque 2, § 4).

On appelle *transposée de la matrice A* la matrice $A^*$

$$A^*=\begin{Vmatrix}\alpha_{11} & \dots & \alpha_{m1}\\ \cdot & \cdot\cdot\cdot & \cdot\\ \alpha_{1n} & \dots & \alpha_{mn}\end{Vmatrix}, \qquad (2)$$

obtenue en échangeant les lignes et les colonnes de $A$.

Les matrices de même dimension, c'est-à-dire composées d'un même nombre de lignes et de colonnes peuvent être additionnées. On appelle *somme* de deux telles matrices $A = \| \alpha_{ij} \|$ et $B = \| \beta_{ij} \|$ la matrice $C = \| \gamma_{ij} \|$ dont les éléments sont égaux à la somme des éléments correspondants de $A$ et $B$: $\gamma_{ij} = \alpha_{ij} + \beta_{ij}$. Ce fait se note:

$$A + B = C.$$

Il est aisé de voir que

$$A + B = B + A,$$
$$(A + B) + C = A + (B + C).$$

On appelle *produit d'un nombre λ par une matrice A* (ou produit d'une matrice $A$ par un nombre $\lambda$) la matrice dont les éléments sont égaux au produit des éléments correspondants de la matrice $A$ par $\lambda$. Donc, $\lambda A = A\lambda$.

EXEMPLE. Soient

$$A=\begin{pmatrix}1&0&2\\2&3&1\end{pmatrix},\quad B=\begin{pmatrix}2&0&1\\2&1&1\end{pmatrix}.$$

Trouver la matrice $\lambda A+\mu B$.

Par définition de la somme des matrices et du produit d'une matrice par un nombre, on a

$$\lambda A=\begin{pmatrix}\lambda&0&2\lambda\\2\lambda&3\lambda&\lambda\end{pmatrix},\quad \mu B=\begin{pmatrix}2\mu&0&\mu\\2\mu&\mu&\mu\end{pmatrix},$$

$$\lambda A+\mu B=\begin{pmatrix}\lambda+2\mu&0&2\lambda+\mu\\2\lambda+2\mu&3\lambda+\mu&\lambda+\mu\end{pmatrix}.$$

## § 4. Système d'équations linéaires. Théorie de Kronecker-Capelli

On appellera *vecteur de dimension n* un système quelconque de $n$ nombres $(x_1, \ldots, x_n)$ et on le désignera par la lettre (grasse) $\boldsymbol{x}$:

$$\boldsymbol{x}=(x_1, \ldots, x_n).$$

Les nombres $x_j$ $(j=1, \ldots, n)$ s'appellent *composantes* du vecteur $\boldsymbol{x}$. Le vecteur

$$\mathbf{0}=(0, \ldots, 0)$$

est le *vecteur nul*.

Soit un système de $n$ équations linéaires à $n$ inconnues

$$\begin{cases} a_{11}x_1+\ldots+a_{1n}x_n=y_1,\\ \ldots\ldots\ldots\ldots\ldots\\ a_{n1}x_1+\ldots+a_{nn}x_n=y_n.\end{cases} \tag{1}$$

Les nombres $a_{kl}$ appelés *coefficients du système* (1) sont donnés. On dira encore que le système (1) est défini par la matrice

$$A=\|a_{kl}\|=\begin{Vmatrix}a_{11}&\ldots&a_{1n}\\ \ldots&\ldots&\ldots\\ a_{n1}&\ldots&a_{nn}\end{Vmatrix}, \tag{2}$$

de ses coefficients.

On s'intéressera à la solubilité du système (1) pour tout vecteur

$$\boldsymbol{y}=(y_1, \ldots, y_n).$$

On dit qu'un vecteur

$$\boldsymbol{x}=(x_1, \ldots, x_n)$$

est *solution du système d'équations* (1) si les nombres $x_j$ vérifient ces équations.

---

[1]) L. Kronecker (1823-1891), mathématicien allemand, A. Capelli (1855-1910), mathématicien italien.

THÉORÈME 1. *Si le déterminant du système* (1) *est différent de zéro:*

$$\Delta = | a_{kl} | \neq 0,$$

*alors le système* (1) *admet une solution unique pour tout vecteur* $y$. *Cette solution est donnée par la formule de Cramer*[1])

$$x_j = \Delta^j / \Delta \quad (j = 1, \ldots, n), \tag{3}$$

*où* $\Delta^j$ *est le déterminant obtenu à partir de* $\Delta$ *en remplaçant les éléments de la j-ième colonne respectivement par* $y_1, \ldots, y_n$:

$$\Delta^j = \begin{vmatrix} a_{11} & \ldots & a_{1,\,j-1} & y_1 & a_{1,\,j+1} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{n,\,j-1} & y_n & a_{n,\,j+1} & \ldots & a_{nn} \end{vmatrix}. \tag{4}$$

*Donc*

$$x_j = \frac{1}{\Delta} \sum_{s=1}^{n} A_{sj} y_s \quad (j = 1, \ldots, n), \tag{3'}$$

*où* $A_{sj}$ *est le cofacteur de l'élément* $a_{sj}$ *dans le déterminant* $\Delta$.

DÉMONSTRATION. Pour trouver le nombre inconnu $x_1$, multiplions la première équation du système (1) par le cofacteur $A_{11}$, la deuxième par $A_{21}, \ldots$, la $n$-ième par $A_{n1}$ et additionnons toutes les équations du système. Comme

$$\sum_{k=1}^{n} a_{k1} x_1 A_{k1} = x_1 \sum_{k=1}^{n} a_{k1} A_{k1} = x_1 \Delta$$

et que

$$\sum_{k=1}^{n} a_{kj} x_j A_{k1} = x_j \sum_{k=1}^{n} a_{kj} A_{k1} = x_j \cdot 0 = 0 \quad (j \neq 1),$$

on obtient $x_1 \Delta = \Delta^1$, où

$$\Delta^1 = \sum_{s=1}^{n} y_s A_{s1}' = \begin{vmatrix} y_1 & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ y_n & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Donc, $x_1 = \Delta^1 / \Delta$, puisque par hypothèse $\Delta \neq 0$.

Dans le cas général, pour $j$ quelconque, multiplions la première équation du système (1) par $A_{1j}$, la deuxième par $A_{2j}, \ldots$, la $n$-ième par $A_{nj}$. En additionnant ces équations on obtient en vertu des propriétés f) et g) l'égalité

$$x_j \sum_{k=1}^{n} a_{kj} A_{kj} = \sum_{k=1}^{n} y_k A_{kj},$$

i.e.

$$x_j \Delta = \Delta^j,$$

[1]) G. Cramer (1704-1752), mathématicien suisse.

où

$$\Delta^j = \sum_{k=1}^{n} y_k A_{kj} = \begin{vmatrix} a_{11} & \dots & a_{1,\,j-1} & y_1 & a_{1,\,j+1} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{n,\,j-1} & y_n & a_{n,\,j+1} & \dots & a_{nn} \end{vmatrix}.$$

D'où l'égalité (3), puisque $\Delta \neq 0$.

On appelle *homogène* un système d'équations de la forme

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0, \\ \dots\dots\dots\dots\dots \\ a_{n1}x_1 + \dots + a_{nn}x_n = 0. \end{cases} \tag{5}$$

Ce système est un cas particulier du système (1) pour $y_1 = \dots$ $\dots = y_n = 0$. Il est évident que le vecteur nul

$$x_1 = 0, \dots, x_n = 0$$

est solution du système homogène (5). Le système homogène (5) peut très bien être vérifié par un vecteur non nul $x = (x_1, \dots, x_n)$, c'est-à-dire par un vecteur dont l'une au moins des composantes $x_i \neq 0$. Cette solution s'appelle *solution non triviale du système homogène* (5) et le vecteur nul, *solution triviale du système homogène* (5).

THÉORÈME 2. *Si le déterminant* $\Delta$ *du système homogène* (5) *est différent de zéro, ce système n'admet que la solution triviale pour solution.*

En effet, d'après la propriété d), tous les déterminants $\Delta^j = 0$ (voir (4)), donc d'après les égalités (3), on a $x_j = 0$ $(j = 1, \dots, n)$.

THÉORÈME 3. *Si le système d'équations* (5) *admet une solution non triviale, son déterminant est nécessairement nul.*

En effet, si $\Delta \neq 0$, le système (5) n'admettrait que la solution triviale en vertu du théorème 2.

En étudiant plus haut le système d'équations linéaires (1) dans le cas où $\Delta \neq 0$, on a montré (théorème 1) que ce système admettait pour tout second membre $y = (y_1, \dots, y_n)$ une solution unique donnée par la formule (3).

Etudions maintenant le système (1) dans le cas où $\Delta = 0$. On admettra que l'un au moins des éléments de la matrice $A$ (voir (2)) n'est pas nul et on désignera par $k$ le rang de $A$ ($k$ = rang $A$). Donc $1 \leqslant k < n$.

On se propose de prouver les règles suivantes (explicitement formulées et démontrées par Kronecker et Capelli).

Pour résoudre le système (1) dont la matrice $A$ est de rang $k$, on doit connaître le rang de la *matrice élargie*

$$B = \begin{Vmatrix} a_{11} & \dots & a_{1n} & y_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & y_n \end{Vmatrix}$$

déduite de $A$ par adjonction de la colonne

$$\left\| \begin{matrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{matrix} \right\| = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}.$$

1) Si rang $B >$ rang $A = k$, le système (1) n'admet pas de solution. Ce système est *impossible*, c'est-à-dire qu'il n'existe pas de vecteur $x = (x_1, \ldots, x_n)$ vérifiant simultanément toutes les équations (1).

2) Si rang $A$ = rang $B = k$, le système (1) admet des solutions. Pour les trouver, nous devons choisir dans le système (1) $k$ équations dont les coefficients forment une matrice de rang $k$, et résoudre ces $k$ équations. Ce système admettra une infinité de solutions représentables par des formules.

Ceci étant, toute solution du système de $k$ équations est automatiquement solution du système de $n - k$ équations restantes.

Les propositions 1) et 2) épuisent toutes les situations, car rang $B$ ne peut être inférieur à $k$.

En effet, la matrice $A$ engendre par hypothèse un déterminant non nul d'ordre $k$ qui est aussi engendré par la matrice $B$.

Exemple 1. Le système

$$\begin{cases} x + y = 1, \\ x - y = 2 \end{cases}$$

a pour déterminant

$$\Delta = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0.$$

Il admet donc une solution unique donnée par les formules

$$x = \Delta^1/\Delta, \quad y = \Delta^2/\Delta,$$

où

$$\Delta^1 = \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3, \quad \Delta^2 = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1,$$

i.e.

$$x = \frac{-3}{-2} = \frac{3}{2}, \quad y = \frac{1}{-2} = -\frac{1}{2}.$$

Exemple 2. Le système

$$\begin{cases} x + y = 1, \\ x + y = 2 \end{cases} \tag{6}$$

a un déterminant nul. La matrice

$$A=\begin{Vmatrix} 1 & 1 \\ 1 & 1 \end{Vmatrix}$$

est de rang 1. La matrice

$$B=\begin{Vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{Vmatrix}$$

est de rang 2. Comme rang $B >$ rang $A$, le système (6) ne possède pas de solution. Ce qui d'ailleurs était manifeste, étant donné qu'un même nombre ne peut être simultanément égal à 1 et à 2.

EXEMPLE 3. Le système

$$\begin{cases} 2x+2y=2, \\ 3x+3y=3 \end{cases} \tag{7}$$

a un déterminant $\Delta = 0$. La matrice

$$A=\begin{Vmatrix} 2 & 2 \\ 3 & 3 \end{Vmatrix}$$

est de rang 1. La matrice

$$B=\begin{Vmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \end{Vmatrix}$$

est aussi de rang 1. Comme rang $A$ = rang $B$ = 1, on prend une équation

$$2x + 2y = 2. \tag{8}$$

Le coefficient en $y$ étant différent de zéro, cette équation est soluble par rapport à $y$:

$$y=\frac{2-2x}{2}=1-x. \tag{9}$$

La formule (9) donne toutes les solutions de l'équation (8). On peut assigner à $x$ une valeur quelconque ($-\infty < x < \infty$) et calculer la valeur correspondante de $y$ d'après la formule (9). On obtient ainsi un vecteur $(x, y)$ vérifiant l'équation (8). L'ensemble de tous les vecteurs $(x, 1 - x)$, où $x \in \,]-\infty, \infty[$, constitue l'ensemble de toutes les solutions de l'équation (8). Ces solutions vérifient automatiquement la deuxième équation du système (7), puisque rang $A$ = rang $B$. Ce résultat était évident sans la théorie des rangs. Les coefficients des équations du système (7) et les seconds membres étant respectivement proportionnels, toute solution d'une équation est solution de l'autre.

EXEMPLE 4. Le système

$$\begin{cases} x+y+\ z=1, \\ 2x+y+2z=1, \\ x+y+3z=2 \end{cases}$$

a pour déterminant

$$\Delta=\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix}=-2\neq 0.$$

Il admet donc une solution unique donnée par les formules

$$x=\frac{\Delta'}{\Delta}=\frac{-1}{2}\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{vmatrix}=-\frac{1}{2}, \qquad y=\frac{\Delta^2}{\Delta}=\frac{-1}{2}\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{vmatrix}=1,$$

$$z=\frac{\Delta^3}{\Delta}=\frac{-1}{2}\begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{vmatrix}=\frac{1}{2}.$$

EXEMPLE 5. Le système

$$\begin{cases} x+y+\ z=1, \\ x+y+2z=1, \\ x+y+3z=2 \end{cases}$$

a pour déterminant

$$\Delta=\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{vmatrix}=0.$$

La matrice

$$A=\begin{Vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \end{Vmatrix}$$

est de rang 2, puisque $\Delta = 0$, mais il existe un déterminant non nul d'ordre deux engendré par $A$. Par exemple

$$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}=1\neq 0.$$

La matrice

$$B=\begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 3 & 2 \end{Vmatrix}$$

est de rang 3, puisque le déterminant qu'elle engendre

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{vmatrix} = 1 \neq 0.$$

Comme rang $B >$ rang $A$, le système ne possède pas de solution.

Exemple 6. Le système

$$\left\{ \begin{aligned} x + y + z &= 1, \\ x + y + 2z &= 1, \\ 2x + 2y + 4z &= 2 \end{aligned} \right. \qquad (10)$$

a pour déterminant

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{vmatrix} = 0.$$

On s'assure sans peine que les matrices

$$A = \begin{Vmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 2 & 2 & 4 \end{Vmatrix}, \quad B = \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 2 & 2 & 4 & 2 \end{Vmatrix}$$

sont de même rang 2. Choisissons dans le système (10) deux équations telles que la matrice $A'$ formée avec leurs coefficients soit de rang 2. On peut prendre ici les deux premières équations ou la première et la troisième. On est donc conduit à étudier le système

$$\left\{ \begin{aligned} x + y + z &= 1, \\ x + y + 2z &= 1. \end{aligned} \right. \qquad (11)$$

Portons dans les seconds membres de ces équations l'une des variables de telle sorte que les coefficients en les autres inconnues forment une matrice $A''$ de rang 2. Ici on peut porter $x$ ou $y$. Le système non homogène

$$\left\{ \begin{aligned} x + z &= 1 - y, \\ x + 2z &= 1 - y \end{aligned} \right. \qquad (12)$$

a pour déterminant

$$\Delta'' = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \neq 0.$$

Il admet donc une solution unique quel que soit le second membre:

$$x = \frac{1}{\Delta''} \begin{vmatrix} 1-y & 1 \\ 1-y & 2 \end{vmatrix} = 1 - y, \quad z = \frac{1}{\Delta''} \begin{vmatrix} 1 & 1-y \\ 1 & 1-y \end{vmatrix} = 0.$$

Donc, le triplet de nombres $(1 - y, y, 0)$, où $y \in ]-\infty, \infty[$, fournit toutes les solutions du système (12) et automatiquement les solutions de la troisième équation du système (10) (on obtient cette équation en multipliant la deuxième par 2).

THÉORÈME 4. *Si le système* (1) *est compatible, c'est-à-dire s'il admet au moins une solution* $x$, *alors on a nécessairement* rang $B$ = rang $A$.

La proposition 1) (cf. page 383) résulte immédiatement du théorème 4, car si rang $B >$ rang $A$, la condition nécessaire de compatibilité du système (1) n'est pas réalisée.

DÉMONSTRATION DU THÉORÈME 4. Supposons que le système (1) admet une solution et que rang $A = k$. Il nous faut prouver que rang $B = k$. Comme rang $A = k$ par hypothèse, il existe un déterminant non nul d'ordre $k$ engendré par la matrice $A$, donc par la matrice $B$. D'où rang $B \geqslant k$. Reste à montrer que tout déterminant d'ordre $k + 1$ engendré par la matrice $B$ est nul. Si un tel déterminant était composé uniquement des éléments $a_{ij}$, il serait forcément nul, car engendré aussi par la matrice $A$ qui est de rang $k$ par hypothèse. Il faut donc prouver que tout déterminant d'ordre $k + 1$ engendré par la matrice $B$ et dont une colonne est composée des nombres $y_j$ est nul. Sans nuire à la généralité on peut admettre que c'est le déterminant

$$\begin{vmatrix} a_{11} & \ldots & a_{1k} & y_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{k+1,1} & \ldots & a_{k+1,k} & y_{k+1} \end{vmatrix}.$$

Tout cas peut être ramené à celui-ci en renumérotant dûment les équations et les inconnues $x_j$. Le système (1) est compatible par hypothèse, c'est-à-dire qu'il existe un vecteur $x = (x_1, \ldots, x_n)$ vérifiant les équations de ce système. Donc $x$ vérifie en particulier les $k + 1$ premières équations du système renuméroté. Par suite,

$$\left\{\begin{array}{l} a_{11}x_1 + \ldots + a_{1k}x_k + \lambda_1 = 0, \\ \ldots\ldots\ldots\ldots\ldots \\ a_{k+1,1}x_1 + \ldots + a_{k+1,k}x_k + \lambda_{k+1} = 0, \end{array}\right. \tag{13}$$

où

$$\left\{\begin{array}{l} \lambda_1 = a_{1,k+1}x_{k+1} + \ldots + a_{1n}x_n - y_1, \\ \ldots\ldots\ldots\ldots\ldots \\ \lambda_{k+1} = a_{k+1,k+1}x_{k+1} + \ldots + a_{k+1,n}x_n - y_{k+1}. \end{array}\right. \tag{14}$$

Composons le système

$$\left\{\begin{array}{l} a_{11}z_1 + \ldots + a_{1k}z_k + \lambda_1 z_{k+1} = 0, \\ \ldots\ldots\ldots\ldots\ldots \\ a_{k+1,1}z_1 + \ldots + a_{k+1,k}z_k + \lambda_{k+1}z_{k+1} = 0. \end{array}\right. \tag{15}$$

En vertu de (13) et de (14) ce système est vérifié par les nombres $x_1, \ldots, x_k, 1$ dont l'un au moins est non nul. Donc, le déterminant du système homogène (15) est nul (cf. théorème 3), c'est-à-dire que

$$0 = \begin{vmatrix} a_{11} & \ldots & a_{1k} & a_{1,k+1}x_{k+1} + \ldots + a_{1n}x_n - y_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1}x_{k+1} + \ldots + a_{k+1,n}x_n - y_{k+1} \end{vmatrix} =$$

$$= \sum_{s=k+1}^{n} x_s \begin{vmatrix} a_{11} & \dots a_{1k} & a_{1s} \\ \dots & \dots & \dots \\ a_{k+1,1} & \dots a_{k+1,k} & a_{k+1,s} \end{vmatrix} - \begin{vmatrix} a_{11} & \dots a_{1k} & y_1 \\ \dots & \dots & \dots \\ a_{k+1,1} & \dots a_{k+1,k} & y_{k+1} \end{vmatrix} =$$

$$= - \begin{vmatrix} a_{11} & \dots a_{1k} & y_1 \\ \dots & \dots & \dots \\ a_{k+1,1} & \dots a_{k+1,k} & y_{k+1} \end{vmatrix}. \quad (16)$$

puisque les déterminants (d'ordre $k+1$!) figurant dans la somme $\sum_s$ sont nuls (car la matrice $A$ est de rang $k$).

On a donc démontré que tout déterminant d'ordre $k+1$ engendré par la matrice $B$ est nul.

Prouvons maintenant la proposition 2) (voir page 383). Le rang de la matrice $A$ étant égal à $k$, le système (1) renferme $k$ équations dont la matrice des coefficients engendre un déterminant non nul d'ordre $k$. En réindiciant les équations et les inconnues on peut faire en sorte que les $k$ premières équations du système (1)

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = y_1, \\ \dots \dots \dots \dots \\ a_{k1}x_1 + \dots + a_{kn}x_n = y_k \end{cases} \quad (1')$$

aient un déterminant

$$\sigma = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \neq 0.$$

Mettons le système (1′) sous la forme:

$$\begin{cases} a_{11}x_1 + \dots + a_{1k}x_k = y_1 - a_{1,k+1}x_{k+1} - \dots - a_{1n}x_n, \\ \dots \dots \dots \dots \dots \dots \\ a_{k1}x_1 + \dots + a_{kk}x_k = y_k - a_{k,k+1}x_{k+1} - \dots - a_{kn}x_n. \end{cases} \quad (17)$$

Comme le déterminant $\sigma \neq 0$, à tout système de nombres $x_{k+1}, \dots, x_n$ est associé un système unique de nombres $x_1, \dots, x_k$ qui, de toute évidence, peut être mis sous la forme:

$$\begin{cases} x_1 = \dfrac{\sigma^1}{\sigma} = \dfrac{1}{\sigma} \sum_{s=1}^{k} (y_s - a_{s,k+1}x_{k+1} - \dots - a_{sn}x_n) A_{s1}, \\ \dots \dots \dots \dots \dots \dots \\ x_k = \dfrac{\sigma^k}{\sigma} = \dfrac{1}{\sigma} \sum_{s=1}^{k} (y_s - a_{s,k+1}x_{k+1} - \dots - a_{sn}x_n) A_{sk}. \end{cases} \quad (18)$$

où $A_{si}$ sont les cofacteurs des éléments $a_{si}$ du déterminant $\sigma$. Donc, toutes les solutions du système (17) sont données par la formule (18). On peut donner des valeurs quelconques à $x_{k+1}, \dots, x_n$, quant aux $x_1, \dots, x_k$ on les calcule avec les formules (18). On voit donc que le système (17) admet une infinité de solutions.

On se propose de prouver que si rang $A$ = rang $B = k$, toute solution $x$ des $k$ premières équations du système (1) est automatiquement solution des autres équations de ce système. Pour fixer les idées on démontre qu'elle est solution de la $(k+1)$-ième équation. On considère les $(k+1)$ premières équations du système (1) mises sous la forme (13) et on démontre que toute solution $x$

des $k$ premières équations (13) est automatiquement solution de la $(k+1)$-ième équation (13). Soit $x$ un vecteur vérifiant les $k$ premières équations (13). Composons les équations (15) par rapport aux inconnues $z_1, \ldots, z_{k+1}$, où les nombres $\lambda_1, \ldots, \lambda_{k+1}$ sont exprimés en fonction des composantes $x_{k+1}, \ldots, x_n$ de $x$ par les formules (14). Le déterminant du système (15) est nul. Ceci ressort des égalités (16) qu'il faut remonter. Le déterminant de droite est nul par hypothèse. Donc, le système (15) admet une solution non triviale $z_1, z_2, \ldots, z_{k+1}$. Le nombre $z_{k+1} \neq 0$. En effet, si $z_1, \ldots, z_k, 0$ était solution du système (15), les nombres $z_1, \ldots, z_k$ devraient être nuls, puisque $\sigma \neq 0$. Donc, le système $z_1, \ldots, z_{k+1}$ serait trivial. Le système (15) étant homogène, il est vérifié non seulement par les nombres $z_1, \ldots, z_{k+1}$, mais aussi par

$$z_1' = z_1/z_{k+1}, \ldots, z_k' = z_k/z_{k+1}, 1.$$

Par conséquent, les nombres $z_1', \ldots, z_k'$ vérifient le système des $k$ premières équations (13) dont le déterminant $\sigma \neq 0$. Or, nous savons que ce système admet une solution unique $x_1, \ldots, x_k$. Donc,

$$z_1' = x_1, \ldots, z_k' = x_k.$$

Si l'on se reporte à la dernière équation de (15), on constate qu'elle est satisfaite par les nombres $(x_1, \ldots, x_k, 1)$, c'est-à-dire que les nombres $(x_1, \ldots, x_k)$ vérifient la $(k+1)$-ième équation du système (13), et en vertu de (14) le vecteur $x = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)$ est solution de la $(k+1)$-ième équation du système (1). Ceci prouve la proposition 2).

REMARQUE 1. Pour résoudre les systèmes d'équations linéaires on peut se servir de la *méthode d'élimination des inconnues* ou *méthode de Gauss* [1]). Soit donné le système

$$\left\{\begin{array}{l} a_{11}x_1 + \ldots + a_{1n}x_n = b_1, \\ \cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n. \end{array}\right. \qquad (1'')$$

Si l'on multiplie une équation de ce système par une constante et qu'on l'ajoute à une autre équation, on obtient un nouveau système équivalent au premier. La matrice $B'$ du nouveau système se déduit de la matrice $B$ $(B \Rightarrow B')$ par addition à une ligne d'une autre ligne multipliée par une constante.

En pratique, au lieu d'écrire le nouveau système d'équations, on se contente simplement de sa matrice $B'$. En effectuant les opérations indiquées sur les équations du système, ou ce qui revient au même, sur la matrice $B$, on peut soit obtenir la solution du système (1″), soit arriver à un système impossible. Comme ce dernier est équivalent au système (1″), ceci prouve que le système (1″) est impossible.

On cite plus bas des exemples d'application de cette méthode.

L'opération $B \Rightarrow B'$ signifie que $B'$ se déduit de $B$ soit a) par multiplication d'une ligne de $B$ par une constante non nulle, soit b) par addition à une ligne d'une ligne multipliée par une constante non nulle, soit c) par l'application d'un nombre fini d'opérations a) et b), soit d) par permutation des lignes.

[1]) C. F. Gauss (1777-1855), mathématicien allemand.

EXEMPLE 7. Résoudre le système

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 5, \\ \phantom{x_1 + {}} x_2 + 2x_3 + 3x_4 = 1, \\ x_1 \phantom{{} + x_2} + 3x_3 + 4x_4 = 2, \\ x_1 + \phantom{2}x_2 + 5x_3 + 6x_4 = 1. \end{cases}$$

On aurait pu certes d'après le théorème 1 calculer les cinq déterminants d'ordre 4 et trouver $x_1$, $x_2$, $x_3$, $x_4$. Mais cela aurait demandé beaucoup de calculs identiques.

Formons la matrice $B$:

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 1 \\ 1 & 0 & 3 & 4 & 2 \\ 1 & 1 & 5 & 6 & 1 \end{pmatrix},$$

où la dernière colonne est constituée des termes constants. Multiplions la première ligne par (—1) et ajoutons-la à la troisième et à la quatrième. On obtient la matrice

$$B_1' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -2 & 0 & 0 & -3 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

Dans la matrice $B_1'$ les éléments de la troisième ligne, qui sont les coefficients en les inconnues, sont tous nuls sauf un. Echangeons cette ligne avec la deuxième. L'élément non nul se retrouve alors sur la diagonale principale:

$$B_1'' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -2 & 0 & 0 & -3 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

Multiplions la deuxième ligne par (—1) pour simplifier davantage l'écriture:

$$B_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & 2 & -4 \end{pmatrix}.$$

Les transformations ultérieures des matrices sont évidentes:

$$B_1 \Rightarrow B_2 = \begin{pmatrix} 1 & 0 & 3 & 4 & 2 \\ 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 2 & 3 & -1/2 \\ 0 & 0 & 2 & 2 & -5/2 \end{pmatrix} \Rightarrow B_3 =$$

$$=\begin{pmatrix}1&0&3&4&2\\0&2&0&0&3\\0&0&2&3&-1/2\\0&0&0&-1&-2\end{pmatrix}\Rightarrow B_4=\begin{pmatrix}1&0&3&4&2\\0&2&0&0&3\\0&0&2&3&-1/2\\0&0&0&1&2\end{pmatrix}\Rightarrow B_5=$$

$$=\begin{pmatrix}1&0&3&0&-6\\0&2&0&0&3\\0&0&2&0&-13/2\\0&0&0&1&2\end{pmatrix}\Rightarrow B_6=\begin{pmatrix}1&0&0&0&15/4\\0&2&0&0&3\\0&0&2&0&-13/2\\0&0&0&1&2\end{pmatrix},$$

d'où $x_4 = 2$, $x_3 = -13/4$, $x_2 = 3/2$, $x_1 = 15/4$. Pour éviter toute erreur il est recommandé de procéder à une vérification en portant les valeurs trouvées dans le système initial.

Appliquons cette méthode à l'exemple 5:

$$\begin{cases} x_1+x_2+\ x_3=1,\\ x_1+x_2+2x_3=1,\\ x_1+x_2+3x_3=2,\end{cases}$$

$$B=\begin{pmatrix}1&1&1&1\\1&1&2&1\\1&1&3&2\end{pmatrix}\Rightarrow B_1=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&2&1\end{pmatrix}\Rightarrow B_2=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&0&1\end{pmatrix}.$$

Donc, le système initial est équivalent au système suivant:

$$\begin{cases} x_1+\ \ x_2+\ \ x_3=1,\\ 0\cdot x_1+0\cdot x_2+\ \ x_3=0,\\ 0\cdot x_1+0\cdot x_2+0\cdot x_3=1.\end{cases}$$

Dans la dernière ligne, le terme constant est égal à un, alors que les coefficients en les inconnues sont nuls, donc ce système est impossible.

Dans l'exemple 6 enfin

$$B=\begin{pmatrix}1&1&1&1\\1&1&2&1\\2&2&4&2\end{pmatrix}\Rightarrow B_1=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&2&0\end{pmatrix}\Rightarrow B_2=$$

$$=\begin{pmatrix}1&1&1&1\\0&0&1&0\\0&0&0&0\end{pmatrix}\Rightarrow B_3=\begin{pmatrix}1&1&1&1\\0&0&1&0\end{pmatrix}\Rightarrow B_4=\begin{pmatrix}1&1&0&1\\0&0&1&0\end{pmatrix}.$$

D'où $x_3 = 0$, $x_1 + x_2 = 1$, c'est-à-dire que le système admet une infinité de solutions: $x_1 = x_1$, $x_2 = 1 - x_1$, $x_3 = 0$, où $x_1$ est un nombre quelconque $(-\infty < x_1 < \infty)$.

REMARQUE 2. Si l'on ne s'intéresse qu'au rang de la matrice $B$, on peut étendre les opérations $B \Rightarrow B'$ non seulement aux lignes mais aussi aux colonnes. Ceci étant, si au cours de ces transformations on se trouve en présence d'une ligne ou d'une colonne constituée uniquement de zéros, on peut la supprimer de la matrice, autrement dit on peut passer à une matrice d'ordre moindre.

Illustrons ceci sur les exemples suivants.

EXEMPLE 8. Trouver le rang de la matrice

$$B=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 0 & 1 & 0 & 1 & 0 & 1 & 1\\ 1 & 2 & 1 & 2 & 1 & 2 & 1\\ 0 & 0 & 0 & 0 & 0 & 4 & 5\end{pmatrix}.$$

Il est clair que $\operatorname{rang} B \leqslant 4$. Ici $a_{11} = 1 \neq 0$. Multiplions la première ligne par $(-1)$ et ajoutons-la à la troisième ligne. On obtient

$$B_1=\begin{pmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\\ 0 & 1 & 0 & 1 & 0 & 1 & 1\\ 0 & 0 & -2 & -2 & -4 & -4 & -6\\ 0 & 0 & 0 & 0 & 0 & 4 & 5\end{pmatrix}.$$

En multipliant maintenant la première colonne par les nombres correspondants et en l'ajoutant aux autres colonnes, on obtient la matrice

$$B_2=\begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 1 & 0 & 1 & 1\\ 0 & 0 & -2 & -2 & -4 & -4 & -6\\ 0 & 0 & 0 & 0 & 0 & 4 & 5\end{pmatrix}.$$

La deuxième colonne est composée de zéros à l'exception de $a_{22} = 1$. En multipliant cette colonne par $(-1)$ et en l'ajoutant à la 4-ième, 6-ième et 7-ième colonne, on obtient

$$B_3=\begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & -2 & -2 & -4 & -4 & -6\\ 0 & 0 & 0 & 0 & 0 & 4 & 5\end{pmatrix} \Rightarrow B_4 =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 5 \end{pmatrix} \Rightarrow B_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix} \Rightarrow$$

$$\Rightarrow B_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \end{pmatrix} \Rightarrow B_7 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Le déterminant d'ordre quatre de la matrice $B_7$ n'est pas nul, donc

$$\text{rang } B = \text{rang } B_7 = 4.$$

Exemple 9. Trouver le rang de la matrice

$$B = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \Rightarrow B_1 = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{pmatrix} \Rightarrow B_2 =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{pmatrix} \Rightarrow B_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \Rightarrow B_4 =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} \Rightarrow B_5 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \Rightarrow B_6 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

c'est-à-dire que rang $B = 2$.

## § 5. Espace à trois dimensions. Vecteurs. Système de coordonnées cartésiennes [1]

Dans ce paragraphe on se propose d'étudier l'espace à trois dimensions. La notion de vecteur dans l'espace a été étudiée en géométrie élémentaire.

On appelle *vecteur libre* (dans l'espace) un segment orienté $\overrightarrow{AB}$ d'origine $A$ et d'extrémité $B$ qui peut se déplacer parallèlement à lui-même. On admet donc que deux segments orientés $\overrightarrow{AB}$ et $\overrightarrow{A_1B_1}$ de même longueur ($|AB| = |A_1B_1|$) et de même sens définissent un même vecteur libre $\boldsymbol{a}$, ce qu'on note $\boldsymbol{a} = \overrightarrow{AB} = \overrightarrow{A_1B_1}$ (fig. 111).

[1] Dans cet ouvrage on étudie d'abord le produit scalaire de vecteurs, ensuite la géométrie analytique de la droite et du plan et après, dans les §§ 11, 12, 13, les notions de produit vectoriel et de produit mixte des vecteurs. On aurait pu traiter ces paragraphes immédiatement après le § 6.

On appelle *longueur* $|\overrightarrow{AB}|$ (ou *module* ou *intensité*) *d'un vecteur* $\overrightarrow{AB}$ le nombre (positif) égal à la longueur du segment $AB$. On écrira aussi $|\overrightarrow{AB}| = |AB|$.

Les vecteurs portés par une même droite ou par des droites paral-

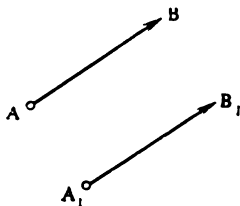



Fig. 111

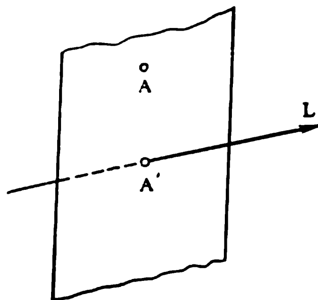



Fig. 112

lèles sont dits *colinéaires*.

Si les points $A$ et $B$ sont confondus, alors $\overrightarrow{AB} = \overrightarrow{AA} = 0$ est considéré comme un vecteur: le *vecteur nul*. Sa longueur est égale à zéro ($|0| = 0$) et son sens n'est pas défini.

On appelle *projection d'un point A sur une droite L* (fig. 112)

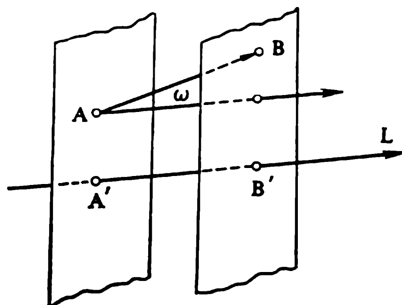



Fig. 113

le point $A'$ en lequel la droite $L$ rencontre le plan perpendiculaire à $L$ passant par $A$.

Soient une droite orientée $L$ (fig. 113) et un vecteur $\boldsymbol{a} = \overrightarrow{AB}$. On appelle *projection du vecteur* $\boldsymbol{a} = \overrightarrow{AB}$ sur la droite orientée $L$ le vecteur $\overrightarrow{A'B'}$, où $A'$ et $B'$ sont les projections respectives des points

$A$ et $B$ sur $L$. Les projections $\overrightarrow{A'B'}$ de vecteurs $\overrightarrow{AB}$ quelconques sur une droite orientée donnée $L$ sont portées par $L$ et orientées dans le même sens ou dans le sens contraire de $L$. Ceci permet d'exprimer la projection des vecteurs sur une droite orientée par des nombres. D'où la deuxième définition.

*On appelle projection d'un vecteur* $\boldsymbol{a} = \overrightarrow{AB}$ *sur une droite orientée* $L$ *le produit de la longueur du vecteur* $\boldsymbol{a} = \overrightarrow{AB}$ *par le cosinus de l'angle* $\omega$ *du vecteur* $\boldsymbol{a}$ *et de la direction de* $L$:

$$\text{pr}_L\, \boldsymbol{a} = |\boldsymbol{a}| \cos(\boldsymbol{a}, L) = |\boldsymbol{a}| \cos\omega \quad (0 \leqslant \omega \leqslant \pi). \tag{1}$$

Le passage de la première définition à la deuxième s'opère comme suit: $\text{pr}_L\, \boldsymbol{a}$ est un nombre égal à la longueur $|\overrightarrow{A'B'}|$ prise avec le

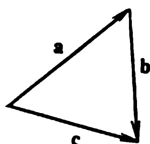


Fig. 114

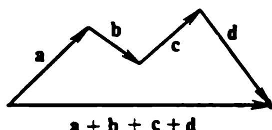


Fig. 115

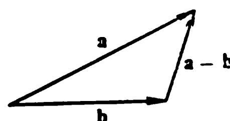


Fig. 116

signe + ou — selon que le vecteur $\overrightarrow{A'B'}$ est de même sens ou de sens contraire à $L$.

En géométrie, on étudie la somme et la différence des vecteurs ainsi que leur produit par des nombres réels. Par définition, le produit $\boldsymbol{a}\alpha = \alpha\boldsymbol{a}$ d'un vecteur $\boldsymbol{a}$ par un nombre $\alpha$ ou d'un nombre $\alpha$ par un vecteur $\boldsymbol{a}$ est un vecteur dont la longueur est égale à $|\alpha\boldsymbol{a}| = |\alpha|\cdot|\boldsymbol{a}|$ et dont le sens est celui de $\boldsymbol{a}$ pour $\alpha > 0$ et le sens contraire pour $\alpha < 0$. Pour $\alpha = 0$ la longueur $|\alpha\boldsymbol{a}|$ est nulle et le vecteur $\alpha\boldsymbol{a}$ se transforme en un vecteur nul (un point) sans sens.

Les figures 114 et 115 nous montrent comment on additionne deux et plusieurs vecteurs, la figure 116, comment on fait la différence de deux vecteurs.

Les projections des vecteurs $\boldsymbol{a}$, $\boldsymbol{b}$ sur une direction donnée $L$ possèdent les propriétés suivantes:

$$\text{pr}_L\, \boldsymbol{a} + \text{pr}_L\, \boldsymbol{b} = \text{pr}_L\, (\boldsymbol{a} + \boldsymbol{b}), \tag{2}$$

$$\text{pr}_L\, (\alpha\boldsymbol{a}) = \alpha\, \text{pr}_L\, \boldsymbol{a}. \tag{3}$$

La propriété (2) se démontre comme suit (voir fig. 117):

$$\text{pr}_L\, \boldsymbol{a} + \text{pr}_L\, \boldsymbol{b} = \overrightarrow{A'B'} + \overrightarrow{B'C'} = \overrightarrow{A'C'} = \text{pr}_L\, \boldsymbol{c} = \text{pr}_L\, (\boldsymbol{a} + \boldsymbol{b}).$$

Comme $\boldsymbol{a} = (\boldsymbol{a} - \boldsymbol{b}) + \boldsymbol{b}$ (voir fig. 116), il vient

$$\text{pr}_L (\boldsymbol{a} - \boldsymbol{b}) + \text{pr}_L \boldsymbol{b} = \text{pr}_L \boldsymbol{a},$$

et par suite

$$\text{pr}_L \boldsymbol{a} - \text{pr}_L \boldsymbol{b} = \text{pr}_L (\boldsymbol{a} - \boldsymbol{b}). \qquad (2')$$

Prouvons (3). Si l'on désigne par $\omega$ l'angle du vecteur $\boldsymbol{a}$ et de la direction $L$, on obtient

pour $\alpha > 0$: $\text{pr}_L (\alpha\boldsymbol{a}) = |\alpha\boldsymbol{a}| \cos \omega = \alpha\, |\boldsymbol{a}| \cos \omega = \alpha\, \text{pr}_L \boldsymbol{a}$;

pour $\alpha < 0$: $\text{pr}_L (\alpha\boldsymbol{a}) = |\alpha\boldsymbol{a}| \cos (\pi - \omega) =$

$$= -\alpha\, |\boldsymbol{a}| \cos (\pi - \omega) = \alpha\, |\boldsymbol{a}| \cos \omega = \alpha\, \text{pr}_L \boldsymbol{a}.$$

En effet, pour $\alpha < 0$ le vecteur $\alpha\boldsymbol{a}$ est orienté dans le sens contraire

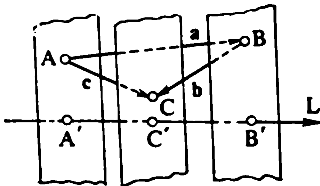



Fig. 117

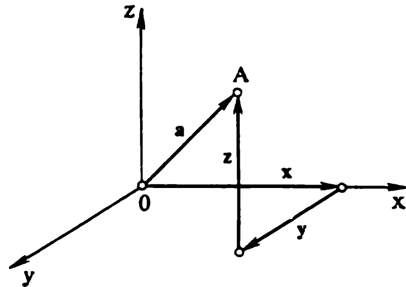



Fig. 118

de $\boldsymbol{a}$ et si $\boldsymbol{a}$ fait un angle $\omega$ avec $L$, alors $\alpha\boldsymbol{a}$ fait un angle de $\pi - \omega$ avec $L$.

Pour $\alpha = 0$ les deux membres de (3) sont nuls.

Appelons *produit scalaire de deux vecteurs* $\boldsymbol{a}$ et $\boldsymbol{b}$ le nombre $(\boldsymbol{a}, \boldsymbol{b})$ égal au produit des longueurs de ces vecteurs par le cosinus de leur angle :

$$\boldsymbol{ab} = (\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}|\,|\boldsymbol{b}| \cos (\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}|\,|\boldsymbol{b}| \cos \omega. \qquad (4)$$

Si l'on se sert de la deuxième définition de la projection d'un vecteur, on peut dire que le *produit scalaire de deux vecteurs* $\boldsymbol{a}$ *et* $\boldsymbol{b}$ *est le produit de la longueur* $|\boldsymbol{b}|$ *par la projection du vecteur* $\boldsymbol{a}$ *sur la direction du vecteur* $\boldsymbol{b}$, *ou de la longueur* $|\boldsymbol{a}|$ *par la projection de* $\boldsymbol{b}$ *sur la direction de* $\boldsymbol{a}$:

$$\boldsymbol{ab} = (\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{a}|\, \text{pr}_{\boldsymbol{a}} \boldsymbol{b} = |\boldsymbol{b}|\, \text{pr}_{\boldsymbol{b}} \boldsymbol{a}.$$

Le produit scalaire est doué des propriétés suivantes :

$$(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{b}, \boldsymbol{a}), \qquad (5)$$

$$(\boldsymbol{a}, \boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a}, \boldsymbol{b}) + (\boldsymbol{a}, \boldsymbol{c}), \qquad (6)$$

$$(\boldsymbol{a}, \alpha\boldsymbol{b}) = \alpha (\boldsymbol{a}, \boldsymbol{b}). \qquad (7)$$

L'égalité (5) résulte immédiatement de la définition du produit scalaire.

L'égalité (6) se prouve comme suit:

$$(\boldsymbol{a},\ \boldsymbol{b}+\boldsymbol{c}) = |\,\boldsymbol{a}\,|\ \mathrm{pr}_{\boldsymbol{a}}(\boldsymbol{b}+\boldsymbol{c}) = |\,\boldsymbol{a}\,|\ \mathrm{pr}_{\boldsymbol{a}}\,\boldsymbol{b} + |\,\boldsymbol{a}\,|\ \mathrm{pr}_{\boldsymbol{a}}\,\boldsymbol{c} = (\boldsymbol{a},\ \boldsymbol{b}) + (\boldsymbol{a},\ \boldsymbol{c}).$$

L'égalité (7) se démontre de la manière suivante:

$$(\boldsymbol{a},\ \alpha\boldsymbol{b}) = |\,\boldsymbol{a}\,|\ \mathrm{pr}_{\boldsymbol{a}}(\alpha\boldsymbol{b}) = |\,\boldsymbol{a}\,|\,\alpha\ \mathrm{pr}_{\boldsymbol{a}}\,\boldsymbol{b} = \alpha\,(\boldsymbol{a},\ \boldsymbol{b}).$$

De (6) et de (7) et compte tenu de (5) il vient

$$(\boldsymbol{a}+\boldsymbol{b},\ \boldsymbol{c}) = (\boldsymbol{a},\ \boldsymbol{c}) + (\boldsymbol{b},\ \boldsymbol{c}). \qquad (6')$$

$$(\alpha\boldsymbol{a},\ \boldsymbol{b}) = \alpha\,(\boldsymbol{a},\ \boldsymbol{b}). \qquad (7')$$

Passons maintenant à la description analytique des vecteurs et des points de l'espace à l'aide des nombres. Introduisons un *système de coordonnées rectangulaires* ($Ox$, $Oy$, $Oz$), c'est-à-dire trois directions deux à deux orthogonales passant par un point $O$, appelées *axes des coordonnées* $x$, $y$, $z$ (fig. 118). Le point $O$ s'appelle *origine des coordonnées*. On admet que pour ce système de coordonnées on a fait choix d'un segment unité pour mesurer les autres segments.

Soit $A$ un point quelconque de l'espace. Le segment orienté $\overrightarrow{OA}$ s'appelle *rayon vecteur* du point $A$. Le rayon vecteur de $A$ définit à son tour un vecteur libre $\boldsymbol{a}$ ($\boldsymbol{a} = \overrightarrow{OA}$). Désignons par $x$, $y$, $z$ les projections de $\boldsymbol{a}$ respectivement sur les axes $Ox$, $Oy$, $Oz$. Ces projections sont les coordonnées du point $A$ : $x$ est l'*abscisse*, $y$, l'*ordonnée*, $z$, la *cote*.

Il existe une correspondance biunivoque entre les points $A$ de l'espace et leurs rayons vecteurs $\overrightarrow{OA}$ ou, ce qui revient au même, les triplets de nombres ($x$, $y$, $z$) qui désignent les coordonnées de $A$ ou les projections de $\overrightarrow{OA}$ sur les axes. Donc, aucune confusion n'est à craindre si l'on appelle le triplet de nombres ($x$, $y$, $z$) point $A$ admettant ces nombres pour coordonnées ou bien vecteur $\boldsymbol{a}$ possédant ces nombres pour projections.

De la définition d'un vecteur libre il s'ensuit que deux vecteurs $\boldsymbol{a} = (x_1,\ y_1,\ z_1)$ et $\boldsymbol{b} = (x_2,\ y_2,\ z_2)$ sont égaux si et seulement si l'on a simultanément $x_1 = x_2$, $y_1 = y_2$, $z_1 = z_2$. On a les égalités

$$(x,\ y,\ z) \pm (x',\ y',\ z') = (x \pm x',\ y \pm y',\ z \pm z'), \qquad (8)$$

$$\alpha\,(x,\ y,\ z) = (\alpha x,\ \alpha y,\ \alpha z). \qquad (9)$$

L'égalité (8) résulte de ce que la projection de la somme ou de la différence de vecteurs (sur l'axe $Ox$ ou $Oy$ ou $Oz$) est égale à la somme ou à la différence des projections des divers termes.

L'égalité (9) résulte, quant à elle, du fait que la projection du vecteur $\alpha\boldsymbol{a}$ (sur l'axe $Ox$ ou $Oy$ ou $Oz$) est égale au produit de $\alpha$ par la projection de $\boldsymbol{a}$.

Désignons par $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ les *vecteurs unités* (de longueur 1) des axes $Ox$, $Oy$, $Oz$. Tout vecteur $(x, y, z)$ peut être représenté sous la forme

$$(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}. \tag{10}$$

En effet

$$\boldsymbol{i} = (1, 0, 0), \quad \boldsymbol{j} = (0, 1, 0), \quad \boldsymbol{k} = (0, 0, 1).$$

Donc, en vertu de (8) et de (9), on a

$$x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) =$$
$$= (x, 0, 0) + (0, y, 0) + (0, 0, z) = (x, y, z).$$

Signalons que

$$\boldsymbol{ii} = \boldsymbol{jj} = \boldsymbol{kk} = 1, \quad \boldsymbol{ij} = \boldsymbol{ik} = \boldsymbol{jk} = 0.$$

Soient maintenant $\boldsymbol{a} = (x, y, z)$, $\boldsymbol{b} = (x', y', z')$. Alors

$$\boldsymbol{ab} = (\boldsymbol{a}, \boldsymbol{b}) = xx' + yy' + zz'. \tag{11}$$

En effet, d'après (6), (7), (6'), (7'), on a

$$\boldsymbol{ab} = (x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})(x'\boldsymbol{i} + y'\boldsymbol{j} + z'\boldsymbol{k}) =$$
$$= xx'\boldsymbol{ii} + xy'\boldsymbol{ij} + xz'\boldsymbol{ik} + yx'\boldsymbol{ji} +$$
$$+ yy'\boldsymbol{jj} + yz'\boldsymbol{jk} + zx'\boldsymbol{ki} + zy'\boldsymbol{kj} + zz'\boldsymbol{kk} =$$
$$= xx' + yy' + zz'.$$

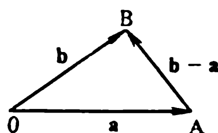



Fig. 119

En particulier, en faisant $\boldsymbol{b} = \boldsymbol{a}$, on obtient

$$|\boldsymbol{a}|^2 = \boldsymbol{aa} = x^2 + y^2 + z^2,$$

d'où

$$|\boldsymbol{a}| = \sqrt{x^2 + y^2 + z^2}.$$

Donc, la distance des points $A = (x, y, z)$ et $B = (x', y', z')$ est égale à (fig. 119):

$$|AB| = |\boldsymbol{b} - \boldsymbol{a}| = +\sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2}.$$

Pour résumer ce qui vient d'être dit modifions légèrement la notation des coordonnées. Plus exactement, soit un système de coordonnées rectangulaires $(Ox_1, Ox_2, Ox_3)$ dans l'espace. Tout point de l'espace se représente par le triplet de nombres

$$\boldsymbol{x} = (x_1, x_2, x_3).$$

Ce point sera désigné par une lettre grasse et appelé aussi vecteur $\boldsymbol{x}$ de composantes $x_1$, $x_2$, $x_3$.

Nous avons montré que dans le langage des triplets $(x_1, x_2, x_3)$ la somme, la différence et le produit par un nombre des vecteurs

s'exprimaient de la manière suivante:

$$\left.\begin{aligned} &(x_1,\ x_2,\ x_3) \pm (x'_1,\ x'_2,\ x'_3) = \\ &\qquad = (x_1 \pm x'_1,\ x_2 \pm x'_2,\ x_3 \pm x'_3), \\ &\alpha (x_1,\ x_2,\ x_3) = (\alpha x_1,\ \alpha x_2,\ \alpha x_3). \end{aligned}\right\} \tag{12}$$

Le produit scalaire des vecteurs $\boldsymbol{x} = (x_1,\ x_2,\ x_3)$ et $\boldsymbol{x}' = (x'_1,\ x'_2,\ x'_3)$ s'exprime en fonction des coordonnées des vecteurs $\boldsymbol{x}$ et $\boldsymbol{x}'$ par la formule

$$\boldsymbol{x}\boldsymbol{x}' = (\boldsymbol{x},\ \boldsymbol{x}') = x_1x'_1 + x_2x'_2 + x_3x'_3. \tag{13}$$

La longueur du vecteur $\boldsymbol{x} = (x_1,\ x_2,\ x_3)$ est le nombre positif

$$|\boldsymbol{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}. \tag{14}$$

Enfin la distance des points $\boldsymbol{x} = (x_1,\ x_2,\ x_3)$ et $\boldsymbol{x}' = (x'_1,\ x'_2,\ x'_3)$ est égale à

$$|\boldsymbol{x} - \boldsymbol{x}'| = \sqrt{(x'_1 - x_1)^2 + (x'_2 - x_2)^2 + (x'_3 - x_3)^2}. \tag{15}$$

L'espace que nous venons d'étudier s'appelle *espace à trois dimensions*, car chacun de ses points se représente par un triplet de nombres réels. On le désignera par $R_3$.

Il semble naturel de désigner un plan par $R_2$. On peut rapporter $R_2$ à un système de coordonnées rectangulaires $(Ox_1,\ Ox_2)$ dans lequel tout point de $R_2$ ou son rayon vecteur se représente par un couple de nombres $(x_1,\ x_2)$. La somme, la différence et le produit par un nombre des vecteurs de $R_2$, le produit scalaire, la longueur d'un vecteur et enfin la distance de deux points sont donnés *mutatis mutandis* par (12), (13), (14) et (15) respectivement.

Les espaces $R_2$ et $R_3$ se généralisent à un espace $R_n$, où $n$ est un entier naturel quelconque.

L'espace $R_n$, $n > 3$, est une invention de l'esprit, une invention géniale du reste qui aide les mathématiciens à comprendre des phénomènes complexes.

Exercices

1. Trouver la longueur des vecteurs (1, 1, 1), (1, —1, 1), (1, 2, 3).
2. Trouver l'angle des vecteurs (1, 0, 1), (1, 2, 2).
3. Soit donné un cube unité (i.e. dont l'arête est de longueur 1). Trouver l'angle: a) d'une diagonale principale et d'une diagonale d'une face; b) de deux diagonales de deux faces.

## § 6. Espace euclidien à $n$ dimensions. Produit scalaire

L'ensemble de tous les systèmes $(x_1,\ \ldots,\ x_n)$ de nombres réels (resp. complexes) s'appelle *espace réel* (resp. complexe) *à n dimensions* et se note $R_n$. On désignera chaque système par une seule lettre (grasse) sans indice:

$$\boldsymbol{x} = (x_1,\ \ldots,\ x_n)$$

et on l'appellera *point* ou *vecteur* de $R_n$. Les nombres $x_1, \ldots, x_n$ s'appellent *coordonnées* du point (vecteur) $x$ ou encore *composantes* du vecteur $x$.

Deux points

$$x = (x_1, \ldots, x_n), \quad x' = (x'_1, \ldots, x'_n)$$

sont égaux si leurs coordonnées respectives sont égales

$$x_j = x'_j \quad (j = 1, \ldots, n).$$

Les systèmes (vecteurs) $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ peuvent être additionnés, retranchés et multipliés par des nombres $\alpha$, $\beta$, ... réels ou complexes selon que $R_n$ est un espace réel ou complexe [1]).

Par définition, on appelle *somme des vecteurs* $x$ et $y$ le vecteur

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \tag{1}$$

*différence*, le vecteur

$$x - y = (x_1 - y_1, \ldots, x_n - y_n). \tag{2}$$

On appelle *produit d'un nombre* $\alpha$ *par un vecteur* $x$ ou d'un vecteur $x$ par un nombre $\alpha$ le vecteur

$$\alpha x = x\alpha = (\alpha x_1, \ldots, \alpha x_n).$$

On a de toute évidence

1) $x + y = y + x$,
2) $(x + y) + z = x + (y + z)$,
3) $x - y = x + (-1)\, y$,
4) $\alpha x + \alpha y = \alpha\, (x + y)$,
5) $\alpha x + \beta x = (\alpha + \beta)\, x$,
6) $\alpha\, (\beta x) = (\alpha\beta)\, x$,
7) $1 \cdot x = x$,

où $\alpha$, $\beta$ sont des nombres et $x, y \in R_n$.

Le nombre (positif)

$$|x| = \sqrt{\sum_{k=1}^{n} |x_k|^2} = \sqrt{\sum_{k=1}^{n} x_k \bar{x}_k} \tag{3}$$

s'appelle *longueur* ou *norme du vecteur* $x = (x_1, \ldots, x_n)$.

La *distance de deux points* $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ est donnée par la formule

$$|x - y| = \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2}. \tag{4}$$

---

[1]) Les nombres complexes sont exposés au § 3, chap. 5.

On appelle *produit scalaire de deux vecteurs* $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ le nombre

$$(x, y) = \sum_{j=1}^{n} x_j \bar{y}_j, \tag{5}$$

où $\bar{y}_j$ est le conjugué complexe de $y_j$ [1]).

Le produit scalaire possède les propriétés suivantes:

a) $(x, x) \geqslant 0$. Ce produit est nul si et seulement si $x = \mathbf{0} = (0, 0, \ldots, 0)$,

b) $\overline{(x, y)} = (y, x)$,

c) $(\alpha x + \beta y, z) = \alpha (x, z) + \beta (y, z)$.

En effet

$$(x, x) = \sum_{j=1}^{n} x_j \bar{x}_j = \sum_{j=1}^{n} |x_j|^2 \geqslant 0,$$

et l'égalité ne peut être réalisée que pour $x_1 = \ldots = x_n = 0$. D'autre part

$$\overline{(x, y)} = \overline{\left(\sum_{j=1}^{n} x_j \bar{y}_j\right)} = \sum_{j=1}^{n} \bar{x}_j y_j = \sum_{j=1}^{n} y_j \bar{x}_j = (y, x).$$

On s'est servi des propriétés de la conjugaison: $\overline{z + z_1} = \bar{z} + \bar{z}_1$ et $\overline{zz_1} = \bar{z} \cdot \bar{z}_1$. Enfin

$$(\alpha x + \beta y, z) = \sum_{j=1}^{n} (\alpha x_j + \beta y_j) \bar{z}_j =$$

$$= \alpha \sum_{j=1}^{n} x_j \bar{z}_j + \beta \sum_{j=1}^{n} y_j \bar{z}_j = \alpha (x, z) + \beta (y, z).$$

Dans la suite on se bornera à l'étude d'un espace réel à $n$ dimensions. Dans cet espace les vecteurs sont multipliés par des nombres réels. On omettra parfois le terme « réel ». Dans un espace réel on peut remplacer le symbole du module (des nombres) par des parenthèses dans les formules (3) et (4):

$$|x| = \sqrt{\sum_{j=1}^{n} x_j^2}, \quad (3') \qquad |x - y| = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2}, \quad (4')$$

le produit scalaire devient

$$(x, y) = \sum_{j=1}^{n} x_j y_j, \tag{5'}$$

car $\bar{y}_j = y_j$, et $(x, y) = (y, x)$, autrement dit, *le produit scalaire est commutatif dans un espace réel.*

---

[1]) Un espace $R_n$ muni du produit scalaire est dit *euclidien*.

Les propriétés a), b), c) du produit scalaire entraînent l'importante *inégalité de Bouniakovski* [1])

$$|(x, y)| \leqslant \sqrt{(x, x)}\sqrt{(y, y)}. \qquad (6)$$

En effet, pour tout nombre réel $\lambda$ on a

$$0 \leqslant (x + \lambda y, x + \lambda y) =$$
$$= (x, x) + \lambda (y, x) + \lambda (x, y) + \lambda^2 (y, y) =$$
$$= (x, x) + 2 (x, y) \lambda + (y, y) \lambda^2 = a + 2b\lambda + c\lambda^2,$$

où $a = (x, x)$, $b = (x, y)$, $c = (y, y)$. On voit que

$$a + 2b\lambda + c\lambda^2 \geqslant 0 \quad \text{pour tout} \quad \lambda \in ]-\infty, \infty[.$$

Donc, le graphe de cette fonction est situé entièrement au-dessus de l'axe des $\lambda$. Or, ceci ne peut avoir lieu que si le discriminant du polynôme est négatif, c'est-à-dire que $b^2 - ac \leqslant 0$, d'où $b^2 \leqslant ac$, et on obtient l'inégalité de Bouniakovski (6).

Dans le langage des composantes des vecteurs $x$ et $y$ l'inégalité (6) s'écrit

$$\left|\sum_{j=1}^{n} x_j y_j\right| \leqslant \sqrt{\sum_{j=1}^{n} x_j^2}\ \sqrt{\sum_{j=1}^{n} y_j^2}. \qquad (7)$$

Donc, l'inégalité (7) est réalisée quels que soient les nombres réels $x_j$, $y_j$.

En vertu de (3) l'inégalité de Bouniakovski peut s'écrire

$$|(x, y)| \leqslant |x|\,|y|.$$

Il existe alors un nombre $\lambda \in [-1, 1]$ unique tel que

$$(x, y) = \lambda\,|x|\,|y|.$$

A noter que sur l'intervalle $[0, \pi]$ la fonction $\cos t$ possède une fonction inverse strictement décroissante définie sur $[-1, 1]$. Donc, pour chaque $\lambda \in [-1, 1]$ il existe un angle unique $\omega \in [0, \pi]$ tel que $\lambda = \cos \omega$. Ainsi on a démontré l'égalité

$$(x, y) = |x|\,|y| \cos \omega. \qquad (8)$$

Le nombre $\omega$ s'appelle *angle des vecteurs à $n$ dimensions* $x$ et $y$, bien que pour $n > 3$ les vecteurs $x$ et $y$ ne soient pas des segments mais des êtres mathématiques abstraits.

Deux vecteurs $x$ et $y$ sont dits *orthogonaux* si leur produit scalaire est nul:

$$(x, y) = \sum_{j=1}^{n} x_j y_j = 0.$$

---

[1]) V. Bouniakovski (1804-1889), mathématicien russe.

De (8) il résulte qu'*une condition nécessaire et suffisante pour que deux vecteurs $x$ et $y$ non nuls soient orthogonaux est que leur angle* $\omega = \pi/2$.

Signalons l'importante *inégalité de Minkowski* [1])

$$|x+y| \leqslant |x| + |y|, \tag{9}$$

ou en composantes

$$\sqrt{\sum_{j=1}^{n} (x_j+y_j)^2} \leqslant \sqrt{\sum_{j=1}^{n} x_j^2} + \sqrt{\sum_{j=1}^{n} y_j^2}. \tag{10}$$

Cette inégalité s'établit comme suit

$$|x+y|^2 = (x+y, x+y) = (x, x) + 2(x, y) + (y, y).$$

L'inégalité de Bouniakovski (6) donne

$$|x+y|^2 \leqslant (x, x) + 2\sqrt{(x, x)}\sqrt{(y, y)} + (y, y) =$$
$$= (\sqrt{(x, x)} + \sqrt{(y, y)})^2,$$

d'où (9). De (9) il s'ensuit que

$$||x| - |y|| \leqslant |x-y|, \tag{11}$$

puisque

$$|x| = |x-y+y| \leqslant |x-y| + |y|,$$
$$|y| = |y-x+x| \leqslant |x-y| + |x|.$$

EXERCICE. Trouver l'angle des vecteurs (1, 0, 1, 0) et (1, 1, 1, 1).

## § 7. Segment. Division d'un segment dans un rapport donné

On se donne deux points arbitraires $x, y \in R_n$ et on considère l'ensemble des points (vecteurs)

$$z = \lambda x + \mu y \quad (\lambda, \mu \geqslant 0, \ \lambda + \mu = 1). \tag{1}$$

On a

$$z = (1-\mu)x + \mu y = x + \mu(y-x)$$
$$(0 \leqslant \mu \leqslant 1) \tag{2}$$

ou

$$z = y + \lambda(x-y) \quad (0 \leqslant \lambda \leqslant 1). \tag{2'}$$

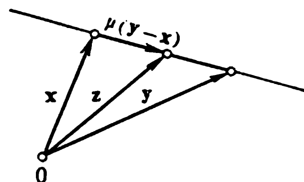


Fig. 120

La relation (2) montre que les points $z$ de $R_3$ couvrent le segment reliant $x$ à $y$. En effet, le rayon vecteur $z$ est la somme du vecteur $x$ et du vecteur $\mu(y-x)$ colinéaire à $y-x$ (fig. 120). Donc, l'ensemble des points (1) est un segment $[x, y]$ de $R_3$. Pour $\mu = 0$, $z = x$;

---

[1]) H. Minkowski (1864-1909), mathématicien balte.

pour $\lambda = 0$ on a $z = y$ ; $z$ est un point quelconque de $[x, y]$ pour tout $\lambda > 0$ $(\mu = 1 - \lambda > 0)$.

On appelle *segment* $[x, y]$ *reliant des points* $x$, $y \in R_n$ l'ensemble de tous les points $z$ de la forme (1). On a le

THÉORÈME. *Le point*

$$z = \lambda x + \mu y \quad (\lambda, \mu \geqslant 0, \quad \lambda + \mu = 1)$$

*partage le segment* $[x, y]$ *de* $R_n$ *en segments dont les longueurs se trouvent dans le rapport* $\mu : \lambda$.

DÉMONSTRATION. De (2) il s'ensuit que $z - x = \mu (y - x)$, donc la distance de points $x$ et $z$ est égale à

$$| z - x | = \mu | y - x |. \tag{3}$$

D'autre part, en vertu de (2′) on a $z - y = \lambda (x - y)$, donc la distance des points $z$ et $y$ est

$$| z - y | = \lambda | x - y |. \tag{4}$$

De (3) et (4) il s'ensuit

$$| z - x | : | z - y | = \mu : \lambda,$$

c.q.f.d.

EXERCICE. Trouver sur le segment $[x, y]$ de $R_n$ le point $z$ qui partage ce segment dans le rapport $p : q$ $(p > 0, q > 0)$.

SOLUTION. Soient

$$\lambda = \frac{q}{p+q}, \qquad \mu = \frac{p}{p+q}.$$

Ces nombres sont positifs et de plus $\lambda + \mu = 1$, $\mu/\lambda = p/q$. Donc, en vertu du théorème ci-dessus le point cherché est

$$z = \lambda x + \mu y = \frac{qx + py}{p+q}. \tag{5}$$

Les coordonnées $z_1, \ldots, z_n$ de $z$ s'expriment en fonction des coordonnées $x_1, \ldots, x_n$ de $x$ et $y_1, \ldots, y_n$ de $y$ par

$$z_j = \frac{qx_j + py_j}{p+q} \quad (j = 1, \ldots, n). \tag{5′}$$

En particulier, le milieu du segment correspond à $p = q = 1$, d'où $\lambda = \mu = 1/2$.

En mécanique on démontre que le point $z$ défini par (5) ou (5′) est le centre de gravité d'un système de points $x$ et $y$ affectés respectivement de masses $q$ et $p$.

A noter que dans $R_3$ l'ensemble des points

$$z = \lambda x + \mu y, \qquad \lambda + \mu = 1,$$

où $\lambda$ et $\mu$ sont de signe arbitraire, est une droite passant par $x$ et $y$. Ceci ressort de la relation (2′).

Dans l'espace $R_n$ ($n > 3$) cet ensemble s'appelle aussi une droite.

EXEMPLE. Trouver les coordonnées du centre de gravité d'un système de points matériels $x^k = (x_1^k, x_2^k, x_3^k)$ de masses respectives $p_k$ ($k = 1, \ldots, M$). Appliquons la formule (5') pour trouver le centre de gravité $z^1$ des points $x^1$ et $x^2$, puis celui des points $z^1$ et $x^3$ (affectés respectivement des masses $p_1 + p_2$ et $p_3$), et ainsi de suite. On obtient en définitive

$$z_j^{M-1} = \frac{1}{p_1 + \cdots + p_M} \sum_{k=1}^{M} p_k x_j^k \quad (j = 1, 2, 3).$$

## § 8. La droite

La notion de droite est une notion primaire en géométrie. Les axiomes de la géométrie nous disent que par deux points il passe une seule droite, et par un point d'une droite on ne peut mener qu'une perpendiculaire à cette droite.

Soient dans le plan un système de coordonnées rectangulaires

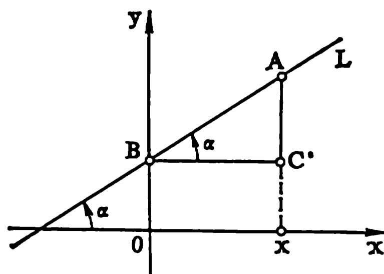



Fig. 121

($Ox$, $Oy$) et une droite $L$ non parallèle à l'axe $Oy$ (fig. 121). On rappelle que l'équation de la droite $L$ est

$$y = kx + l, \tag{1}$$

où $k = \operatorname{tg} \alpha$, $\alpha$ étant l'angle de la droite $L$ et de l'axe $Ox$, $l$ l'ordonnée à l'origine de $L$ ($l = OB$).

Dire que (1) est l'équation de la droite $L$ revient à dire que $L$ est le lieu géométrique des points dont les coordonnées ($x$, $y$) vérifient (1). En effet, pour le point courant $A$ de la droite $L$ on a (voir fig. 121): $BC = x$, $AC = y - l$ et

$$k = \operatorname{tg} \alpha = \frac{y - l}{x}, \tag{1'}$$

d'où (1). Inversement, la relation (1) est équivalente à (1'), laquelle exprime de toute évidence le fait que le point $(x, y)$ est situé sur la droite $L$. Sur la figure 121, l'angle $\alpha$ est aigu. Les raisonnements sont les mêmes pour un angle obtus.

Soit l'équation

$$Ax + By + C = 0, \qquad (2)$$

où $A$, $B$, $C$ sont des nombres donnés et en outre $A$ et $B$ ne sont pas simultanément nuls.

Si $B \neq 0$, l'équation (2) s'écrit

$$y = -\frac{A}{B} x - \frac{C}{B} \qquad (2')$$

ou encore

$$y = kx + l$$

où $k = -\frac{A}{B}$, $l = -\frac{C}{B}$. Les équations (2) et (2') étant équivalentes (en effet, tout point $(x, y)$ vérifiant l'une vérifie l'autre), l'équation (2) pour $B \neq 0$ est l'équation d'une droite de pente $-A/B$ (i.e. $\operatorname{tg} \alpha = -A/B$) et d'ordonnée à l'origine $l = -C/B$. Pour $B = 0$, l'équation (2) devient

$$Ax + C = 0 \qquad (A \neq 0)$$

ou encore

$$x = a \qquad (a = -C/A).$$

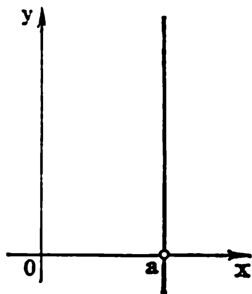


Fig. 122

C'est l'équation d'une droite parallèle à l'axe $Oy$. C'est notamment le lieu géométrique des points $(x, y)$ dont l'abscisse $x$ est égale à une constante $a$. La figure 122 représente une telle droite pour $a > 0$.

En résumé donc, l'équation (2), où $A$, $B$, $C$ sont des nombres donnés et de plus $A$ et $B$ ne sont pas simultanément nuls, est l'équation d'une droite. Pour $B \neq 0$, cette droite n'est pas parallèle à l'axe $Oy$. Pour $A = 0$, elle est parallèle à l'axe $Ox$ $(y = -C/B)$. Pour $B = 0$ elle est parallèle à l'axe $Oy$. Signalons que l'équation de l'axe $Ox$ est $y = 0$ et celle de l'axe $Oy$, $x = 0$.

L'équation (2) est l'*équation générale de la droite*. Toute droite, indépendamment de sa position par rapport au système de coordonnées, peut être représentée par une équation de la forme (2) pour des valeurs adéquates des constantes $A$, $B$, $C$. On rappelle que les *nombres $A$ et $B$ ne sont pas simultanément nuls*. Le nombre $k$ de l'équation (1) s'appelle *coefficient angulaire de la droite*.

Voyons quelques exercices importants.

EXERCICE 1. Former l'équation d'une droite de coefficient angulaire $k$ passant par un point donné $(x_0, y_0)$.

SOLUTION. La droite de coefficient angulaire $k$ a pour équation

$$y = kx + l, \tag{3}$$

où $l$ est un nombre quelconque. Le point $(x_0, y_0)$ appartenant à cette droite, on a

$$y_0 = kx_0 + l. \tag{4}$$

En retranchant (4) de (3) on obtient l'équation cherchée

$$y - y_0 = k(x - x_0). \tag{5}$$

EXERCICE 2. Former l'équation de la droite qui passe par deux points donnés $(x_1, y_1)$ et $(x_2, y_2)$. On suppose que ces points ne sont pas confondus.

SOLUTION. Supposons que $x_1 \neq x_2$. Il est évident que cette droite n'est pas parallèle à l'axe $Oy$ et par conséquent son équation est de la forme

$$y - y_1 = k(x - x_1), \tag{6}$$

où $k$ est un nombre. L'équation (6) exprime que la droite passe par le point $(x_1, y_1)$. Pour qu'elle passe par le point $(x_2, y_2)$ il faut que l'on ait

$$y_2 - y_1 = k(x_2 - x_1). \tag{7}$$

En divisant membre à membre (6) par (7), on obtient

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \tag{8}$$

C'est l'équation de la droite passant par les points $(x_1, y_1)$ et $(x_2, y_2)$.

REMARQUE 1. Si $y_2 - y_1 = 0$, on aurait formellement obtenu la relation

$$\frac{y - y_1}{0} = \frac{x - x_1}{x_2 - x_1}.$$

Même si cette égalité n'a pas de sens, on l'utilise sous cette forme pour la commodité. Si l'on chasse les dénominateurs, on obtient

$$y - y_1 = 0 \cdot \frac{x - x_1}{x_2 - x_1} = 0$$

ou

$$y = y_1. \tag{9}$$

Le cas $x_1 = x_2 = c$ nous conduit à la solution $x = c$.

EXERCICE 3. Trouver l'angle des droites

$$y = k_1 x + l_1, \qquad y = k_2 x + l_2.$$

SOLUTION. On a $k_1 = \operatorname{tg} \alpha_1$, $k_2 = \operatorname{tg} \alpha_2$, où $\alpha_1$ et $\alpha_2$ sont les angles de ces droites avec l'axe $Ox$. On a (voir fig. 123)

$$\operatorname{tg} \omega = \operatorname{tg} (\alpha_1 - \alpha_2) = \frac{\operatorname{tg} \alpha_1 - \operatorname{tg} \alpha_2}{1 + \operatorname{tg} \alpha_1 \operatorname{tg} \alpha_2} = \frac{k_1 - k_2}{1 + k_1 k_2}. \qquad (10)$$

C'est la *formule de l'angle de deux droites.*

La condition $1 + k_1 k_2 = 0$ ou

$$k_1 \cdot k_2 = -1 \qquad (11)$$

exprime l'*orthogonalité des deux droites.* La *condition de parallélisme des droites* ($\operatorname{tg} \omega = 0$) s'écrit

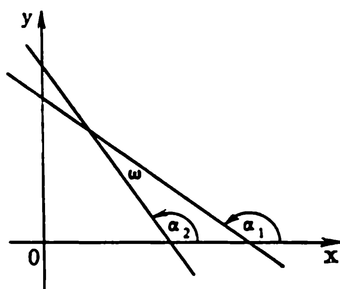



Fig. 123

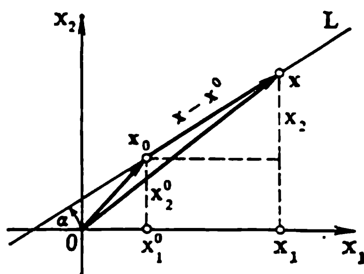



Fig. 124

$$k_1 = k_2. \qquad (12)$$

Ecrivons l'équation (2) sous la forme suivante:

$$A_1 x_1 + A_2 x_2 + C = 0, \qquad (2')$$

où $x = x_1$, $y = x_2$, $A = A_1$, $B = A_2$. Pour $A_1 \neq 0$, $A_2 \neq 0$, $C \neq 0$ l'équation (2') devient:

$$\frac{x_1}{a} + \frac{x_2}{b} = 1, \quad a = \frac{-C}{A_1}, \quad b = \frac{-C}{A_2}. \qquad (13)$$

L'équation (13) s'appelle *équation de la droite en fonction de ses coordonnées à l'origine.* Cette droite coupe l'axe $Ox_1$ (la droite $x_2 = 0$) au point $(a, 0)$ et l'axe $Ox_2$ au point $(0, b)$.

Si une droite vérifiant l'équation (2') passe par le point $(x_1^0, x_2^0)$, on a

$$A_1 x_1^0 + A_2 x_2^0 + C = 0. \qquad (14)$$

En retranchant (14) de (2'), on obtient

$$A_1 (x_1 - x_1^0) + A_2 (x_2 - x_2^0) = 0. \qquad (15)$$

L'équation (15) est l'*équation d'une droite passant par le point* $(x_1^0, x_2^0)$.

Si l'on introduit les vecteurs $A = (A_1, A_2)$, $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{x}^0 = (x_1^0, x_2^0)$, le premier membre de (15) peut être traité comme le produit scalaire du vecteur $A$ par le vecteur $\boldsymbol{x} - \boldsymbol{x}^0$. Donc, l'équation

(15) prend la forme vectorielle

$$A\,(x - x^0) = 0. \tag{15'}$$

Le vecteur $x - x^0$ est porté par la droite $L$ (fig. 124). Donc, la relation (15') exprime que le vecteur $A = (A_1, A_2)$ est orthogonal à $L$. La signification géométrique des coefficients $A_1$ et $A_2$ est maintenant claire.

Soient les droites

$$A_1x_1 + A_2x_2 + C_1 = 0 \quad (L_1), \tag{16}$$

$$B_1x_1 + B_2x_2 + C_2 = 0 \quad (L_2). \tag{17}$$

Les vecteurs $A = (A_1, A_2)$ et $B = (B_1, B_2)$ étant perpendiculaires respectivement aux droites $L_1$ et $L_2$, l'angle $\varphi$ de ces dernières est

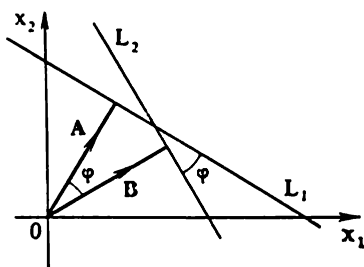



Fig. 125

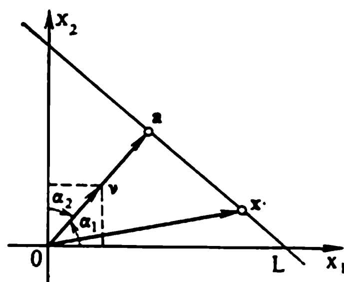



Fig. 126

égal à celui des vecteurs $A$ et $B$ (fig. 125). L'angle $\varphi$ est donné par la formule

$$\cos\varphi = \frac{(A,\,B)}{|A|\,|B|} = \frac{A_1B_1 + A_2B_2}{\sqrt{A_1^2 + A_2^2}\,\sqrt{B_1^2 + B_2^2}}. \tag{18}$$

Remarque 2. Si $\varphi$ est l'angle de deux droites, $\pi - \varphi$ l'est également. Le nombre (18) peut être positif ou négatif. L'une de ces valeurs correspond à $\varphi$, l'autre à $\pi - \varphi$.

De (18) on déduit la *condition d'orthogonalité* de $L_1$ et $L_2$ ($\varphi = \pi/2$, $\cos\varphi = 0$):

$$A_1B_1 + A_2B_2 = 0. \tag{19}$$

Si les droites $L_1$ et $L_2$ sont *parallèles*, les vecteurs $A$ et $B$ sont colinéaires et $A = \lambda B$, où $\lambda$ est un nombre réel. De là on déduit la *condition de parallélisme* de deux droites:

$$\frac{A_1}{B_1} = \frac{A_2}{B_2}. \tag{20}$$

Soient donnés dans un système de coordonnées rectangulaires (fig. 126) une droite $L$ ne passant pas par l'origine et un vecteur $a$

perpendiculaire à $L$ d'origine $O$ et d'extrémité en $L$. Le vecteur $\boldsymbol{a}$ définit entièrement la droite $L$ (il existe en effet une seule droite perpendiculaire à $\boldsymbol{a}$ en son extrémité). Soient $p = |\boldsymbol{a}|$, $\boldsymbol{\nu} = (\cos \alpha_1, \cos \alpha_2)$ le vecteur unité de même sens que $\boldsymbol{a}$. Ici $\alpha_i$ est l'angle de $\boldsymbol{a}$ (ou de $\boldsymbol{\nu}$) et de l'axe $Ox_i$ $(i = 1, 2)$; $\cos^2 \alpha_1 + \cos^2 \alpha_2 = 1$ $(\cos \alpha_2 = \sin \alpha_1)$. Soit $\boldsymbol{x} = (x_1, x_2)$ le point courant de la droite $L$, ou ce qui revient au même, son rayon vecteur. La projection de $\boldsymbol{x}$ sur $\boldsymbol{\nu}$ est visiblement égale à $p$, autrement dit le produit scalaire du rayon vecteur d'un point arbitraire $\boldsymbol{x}$ de la droite $L$ par le vecteur $\boldsymbol{\nu}$ est égal à $p$:

$$(\boldsymbol{x}, \boldsymbol{\nu}) = |\boldsymbol{\nu}| \operatorname{pr}_{\boldsymbol{\nu}} \boldsymbol{x} = p. \tag{21}$$

Ceci est l'équation vectorielle de $L$. En effet, si un point (rayon vecteur) $\boldsymbol{x}$ satisfait à l'équation (21), il appartient à $L$ (la projection sur $\boldsymbol{\nu}$ d'un point n'appartenant pas à $L$ est différente de $p$).

Si la droite $L$ passe par l'origine des coordonnées, son équation est aussi de la forme (21), où $\boldsymbol{\nu}$ est le vecteur unité orthogonal à $L$.

En coordonnées l'équation (21) s'écrit

$$x_1 \cos \alpha_1 + x_2 \cos \alpha_2 = p \quad (p \geqslant 0) \tag{21'}$$

ou

$$x_1 \cos \alpha_1 + x_2 \sin \alpha_1 = p \quad (p \geqslant 0). \tag{21''}$$

L'équation (21′) (ou (21″)) s'appelle *équation de la droite sous la forme normale.*

Si la droite $L$ est définie par l'équation générale (2′)

$$A_1x_1 + A_2x_2 + C = 0,$$

on peut ramener cette équation à la forme normale en la multipliant par

$$M = \pm 1/\sqrt{A_1^2 + A_2^2}, \tag{22}$$

pris avec le signe contraire de $C$ $(p = -MC \geqslant 0)$. Le nombre $M$ s'appelle *facteur de normalisation.* Comme

$$(MA_1)^2 + (MA_2)^2 = 1,$$

il existe un angle $\alpha_1 \in [0, 2\pi[$ unique tel que

$$MA_1 = \cos \alpha_1, \quad MA_2 = \sin \alpha_1. \tag{23}$$

En définitive on obtient l'équation (21), où $p = -MC \geqslant 0$. On rappelle que $p$ est égale à la distance de la droite à l'origine des coordonnées.

Exercice 4. Trouver la distance $d$ du point $\boldsymbol{x}^0 = (x_1^0, x_2^0)$ à la droite $L$ d'équation

$$A_1x_1 + A_2x_2 + C = 0. \tag{24}$$

Solution. Supposons que

$$(\boldsymbol{x}, \boldsymbol{\nu}) - p = 0 \tag{25}$$

est l'équation normale de la droite $L$. Si $C \neq 0$, alors $p$ $(p > 0)$ est la longueur du vecteur $\overrightarrow{OA}$ perpendiculaire à $L$, $\mathbf{v}$ le vecteur unité de même sens que $\overrightarrow{OA}$ $(p = |\overrightarrow{OA}|$, $\mathbf{v} = \overrightarrow{OA}/p$, fig. 127). Soit $x$ un point arbitraire de $L$. Il est évident que pour trouver la distance

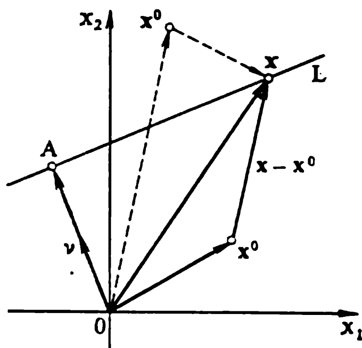



Fig. 127

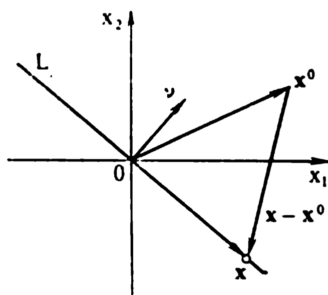



Fig. 128

d'un point $x^0$ à $L$, il faut projeter le vecteur $x - x^0$ sur la direction du vecteur $\mathbf{v}$ et prendre le module de cette projection:

$$d = |\operatorname{pr}_{\mathbf{v}}(x - x^0)| = |(x - x^0, \mathbf{v})| =$$
$$= |(x, \mathbf{v}) - (x^0, \mathbf{v})| = |p - (x^0, \mathbf{v})| = |(x^0, \mathbf{v}) - p|.$$

On obtient la formule

$$d = |(x^0, \mathbf{v}) - p|. \tag{26}$$

Donc, pour trouver la distance $d$, il faut ramener l'équation (24) à la forme normale, faire passer $p$ dans le premier membre, remplacer $x_1$ et $x_2$ respectivement par les coordonnées $x_1^0$ et $x_2^0$ du point $x^0$ et prendre le module de l'expression obtenue.

Dans le langage des coefficients $A_1$, $A_2$, $C$, l'égalité (26) devient:

$$d = \frac{|A_1x_1^0 + A_2x_2^0 + C|}{\sqrt{A_1^2 + A_2^2}}. \tag{26'}$$

La formule (26) et, par conséquent, la formule (26′) sont valables pour $C = 0$. Dans ce cas, $p = 0$ et $\mathbf{v}$ est l'un des deux vecteurs unités orthogonaux à $L$ (fig. 128). Désormais,

$$d = |\operatorname{pr}_{\mathbf{v}}(x - x^0)| = |(\mathbf{v}, x - x^0)| =$$
$$= |(\mathbf{v}, x) - (\mathbf{v}, x^0)| = |(\mathbf{v}, x^0)|$$

ou

$$d = \frac{|A_1x_1^0 + A_2x_2^0|}{\sqrt{A_1^2 + A_2^2}},$$

c'est-à-dire la formule (26′) pour $C = 0$.

REMARQUE 3. Sur la figure 127 on voit que si l'origine $O$ et le point $\boldsymbol{x}^0$ sont situés: a) d'un même côté de $L$, alors l'angle de $\mathbf{v}$ et de $\boldsymbol{x} - \boldsymbol{x}^0$ est aigu et $d = p - (\boldsymbol{x}^0, \mathbf{v})$ ; b) de part et d'autre de $L$, alors l'angle de $\boldsymbol{x} - \boldsymbol{x}^0$ et $\mathbf{v}$ est obtus et $d = (\boldsymbol{x}^0, \mathbf{v}) - p$.

EXERCICE 5. Trouver la distance du point (1, 1) à la droite $2x + \sqrt{5}y - \sqrt{5} = 0$.

## § 9. Equation du plan

Soit un espace $R_3$ rapporté à un système de coordonnées rectangulaires $(Ox_1, Ox_2, Ox_3)$. On considère un vecteur $\boldsymbol{a}$ d'origine $O$ par l'extrémité duquel on mène un plan orthogonal à $\boldsymbol{a}$ (fig. 129).

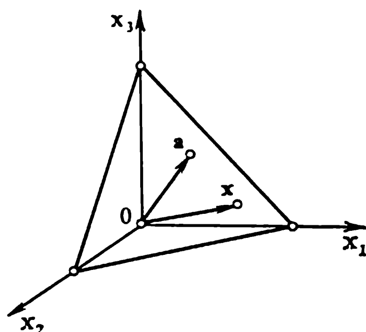



Fig. 129

Désignons par $\boldsymbol{x} = (x_1, x_2, x_3)$ le point courant du plan $L$. La lettre $\boldsymbol{x}$ représente également le rayon vecteur du point $\boldsymbol{x}$.

Soient $p = |\boldsymbol{a}|$ et

$$\mathbf{v} = (\cos\alpha_1, \cos\alpha_2, \cos\alpha_3)$$

un vecteur unité de même sens que $\boldsymbol{a}$. Les angles $\alpha_1, \alpha_2, \alpha_3$ sont les angles formés par le vecteur $\mathbf{v}$ respectivement avec les axes $Ox_1$, $Ox_2$, $Ox_3$. La projection de tout point $\boldsymbol{x} \in L$ sur le vecteur $\mathbf{v}$ est visiblement une quantité constante égale à $p$:

$$(\boldsymbol{x}, \mathbf{v}) = p \quad (p \geqslant 0). \tag{1}$$

L'équation (1) a un sens pour $p = 0$. Dans ce cas, le plan $L$ passe par l'origine des coordonnées $O$ ($\boldsymbol{a} = 0$), et $\mathbf{v}$ est un vecteur unité issu de $O$ et orthogonal à $L$ dans n'importe quel sens, autrement dit le vecteur $\mathbf{v}$ est défini au signe près. L'équation (1) s'appelle *équation du plan L sous la forme vectorielle.* En coordonnées elle s'écrit

$$x_1 \cos\alpha_1 + x_2 \cos\alpha_2 + x_3 \cos\alpha_3 = p \quad (p \geqslant 0) \tag{1'}$$

et s'appelle *équation du plan sous la forme normale.*

Si l'on multiplie l'équation (1′) par un nombre non nul, on obtient l'équation équivalente

$$A_1x_1 + A_2x_2 + A_3x_3 + B = 0, \tag{2}$$

qui définit le même plan. Les nombres $A_1, A_2, A_3$ ne sont pas tous nuls. L'équation (2), où $A_1, A_2, A_3$ ne sont pas simultanément nuls, s'appelle *équation du plan sous la forme générale.*

L'équation (2) peut être ramenée à la forme normale par multiplication par le nombre

$$M = \pm 1/\sqrt{A_1^2 + A_2^2 + A_3^2},$$

pris avec le signe contraire de celui de $B$. Le nombre $p = -MB$ ne sera pas négatif et l'équation (2) se transforme en l'équation équivalente

$$MA_1x_1 + MA_2x_2 + MA_3x_3 = p. \tag{3}$$

Ici

$$(MA_1)^2 + (MA_2)^2 + (MA_3)^2 = 1.$$

Ceci montre que

$$\mathbf{v} = (MA_1, MA_2, MA_3)$$

est un vecteur unité ($|\mathbf{v}| = 1$). Ses projections sur les axes de coordonnées sont respectivement égales à

$$MA_1 = \cos\alpha_1,$$
$$MA_2 = \cos\alpha_2,$$
$$MA_3 = \cos\alpha_3,$$

où $\alpha_1, \alpha_2, \alpha_3$ sont les angles du vecteur $\mathbf{v}$ respectivement avec les axes $Ox_1, Ox_2, Ox_3$. Dans ces notations, l'équation (3) devient

$$x_1\cos\alpha_1 + x_2\cos\alpha_2 + x_3\cos\alpha_3 = p \quad (p \geqslant 0), \tag{3'}$$

c'est-à-dire l'équation du plan (2) sous la forme normale.

L'équation d'un plan étant donnée sous la forme générale (2), si l'on veut savoir comment ce plan est disposé par rapport au système de coordonnées, il suffit de ramener l'équation (2) à la forme normale en la multipliant par le facteur de normalisation $M$.

De l'équation (2), sans calculs supplémentaires, on déduit que:

1) si $B = 0$, le plan passe par l'origine des coordonnées;

2) le vecteur $A = (A_1, A_2, A_3)$ est perpendiculaire au plan, car il est colinéaire au vecteur unité $\mathbf{v} = (MA_1, MA_2, MA_3)$, perpendiculaire à ce plan.

L'équation

$$A_1x_1 + A_2x_2 + B = 0 \tag{4}$$

est un cas particulier de l'équation (2). Dans le plan $(x_1, x_2)$ l'équation (4) définit une droite, dans l'espace $(x_1, x_2, x_3)$, un plan $L$ perpendiculaire au plan de coordonnées $(x_1, x_2)$ et passant par cette droite. Les coordonnées $x_1, x_2$ de tout point $(x_1, x_2, x_3)$ de $L$ vérifient l'équation (4) quel que soit $x_3$. L'équation

$$A_1x_1 + B = 0 \qquad (A_1 \neq 0) \tag{5}$$

est un cas particulier de l'équation (4). On peut la mettre encore sous la forme

$$x_1 = C \qquad (C = -B/A_1). \tag{5'}$$

Dans l'espace $(x_1, x_2, x_3)$, l'équation (5') définit le lieu géométrique des points dont la première coordonnée $x_1$ est une constante $C$. Il est évident que (5') définit un plan parallèle au plan de coordonnées $(x_2, x_3)$ (ou perpendiculaire à l'axe $Ox_1$).

Equation d'un plan en fonction de ses coordonnées à l'origine. Si $A_i \neq 0$ $(i = 1, 2, 3)$ et $B \neq 0$, on peut mettre l'équation (2) sous la forme

$$\frac{x_1}{a_1} + \frac{x_2}{a_2} + \frac{x_3}{a_3} = 1, \tag{6}$$

où $a_i = -B/A_i$ $(i = 1, 2, 3)$. L'équation (6) est l'*équation du plan en fonction de ses coordonnées à l'origine*. Ce plan (fig. 130) coupe l'axe $Ox_1$ en $(a_1, 0, 0)$, l'axe $Ox_2$ en $(0, a_2, 0)$, l'axe $Ox_3$ en $(0, 0, a_3)$. L'équation (6) nous permet d'imaginer sans peine comment est situé le plan par rapport au système de coordonnées.

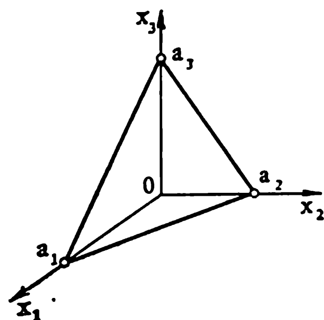


Fig. 130

Equation d'un plan passant par un point. Si un point $x^0 = (x_1^0, x_2^0, x_3^0)$ est situé sur le plan (2), ses coordonnées vérifient l'équation (2):

$$\sum_{i=1}^{3} A_i x_i^0 + B = 0. \tag{7}$$

En retranchant (7) de (2), on obtient

$$\sum_{i=1}^{3} A_i (x_i - x_i^0) = 0. \tag{8}$$

L'équation (8) est l'*équation d'un plan passant par* $x^0 = (x_1^0, x_2^0, x_3^0)$. En écriture vectorielle, l'équation (8) devient

$$A (x - x^0) = 0. \tag{8'}$$

Comme le vecteur $x - x^0$ d'origine $x^0$ appartient au plan (1), la relation (8') nous dit que le vecteur $A$ est orthogonal au plan (1), ce que du reste nous avions établi précédemment par d'autres raisonnements.

Equation d'un plan passant par trois points. Soient donnés trois points

$$x^1 = (x_1^1, x_2^1, x_3^1),$$
$$x^2 = (x_1^2, x_2^2, x_3^2),$$
$$x^3 = (x_1^3, x_2^3, x_3^3)$$

non alignés. On demande l'équation du plan passant par ces points. On sait du cours de géométrie qu'il existe un plan et un seul passant par trois points. Comme ce plan passe par le point $x^1$, son équation est

$$A_1 (x_1 - x_1^1) + A_2 (x_2 - x_2^1) + A_3 (x_3 - x_3^1) = 0, \tag{9}$$

où $A_1$, $A_2$, $A_3$ ne sont pas simultanément nuls. Comme ce plan passe par les points $x^2$ et $x^3$, on a de façon analogue

$$\begin{cases} A_1(x_1^2-x_1^1)+A_2(x_2^2-x_2^1)+A_3(x_3^2-x_3^1)=0, \\ A_1(x_1^3-x_1^1)+A_2(x_2^3-x_2^1)+A_3(x_3^3-x_3^1)=0. \end{cases} \tag{10}$$

Formons le système d'équations linéaires homogènes en $(z_1, z_2, z_3)$:

$$\begin{cases} (x_1-x_1^1)z_1+(x_2-x_2^1)z_2+(x_3-x_3^1)z_3=0, \\ (x_1^2-x_1^1)z_1+(x_2^2-x_2^1)z_2+(x_3^2-x_3^1)z_3=0, \\ (x_1^3-x_1^1)z_1+(x_2^3-x_2^1)z_2+(x_3^3-x_3^1)z_3=0. \end{cases} \tag{11}$$

$x=(x_1, x_2, x_3)$ est un point arbitraire vérifiant l'équation du plan (9). En vertu de (9) et de (10), le système (11) est satisfait par le vecteur non trivial $A=(A_1, A_2, A_3)$, donc le déterminant de ce système est nul:

$$\begin{vmatrix} x_1-x_1^1 & x_2-x_2^1 & x_3-x_3^1 \\ x_1^2-x_1^1 & x_2^2-x_2^1 & x_3^2-x_3^1 \\ x_1^3-x_1^1 & x_2^3-x_2^1 & x_3^3-x_3^1 \end{vmatrix}=0. \tag{12}$$

En développant ce déterminant suivant les éléments de la première ligne, on obtient une équation de la forme (9), c'est-à-dire l'équation d'un plan. En vertu des propriétés du déterminant ce plan passe par les points $x^1$, $x^2$, $x^3$. C.q.f.d.

L'équation (12) se met sous la forme

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ x_1^1 & x_2^1 & x_3^1 & 1 \\ x_1^2 & x_2^2 & x_3^2 & 1 \\ x_1^3 & x_2^3 & x_3^3 & 1 \end{vmatrix}=0. \tag{13}$$

Le déterminant (13) ne change pas si l'on retranche la deuxième ligne de la première, de la troisième et de la quatrième. En développant le déterminant obtenu suivant les éléments de la quatrième colonne, on obtient l'équation (12).

A n g l e  d e  d e u x  p l a n s. Soient les plans

$$\begin{cases} A_1x_1+A_2x_2+A_3x_3+B=0, \\ A_1'x_1+A_2'x_2+A_3'x_3+B'=0. \end{cases} \tag{14}$$

On sait que les vecteurs $A=(A_1, A_2, A_3)$ et $A'=(A_1', A_2', A_3')$ sont orthogonaux respectivement au premier et au second plan, donc l'angle $\varphi$ de $A$ et de $A'$ est égal au rectiligne du dièdre formé par ces deux plans. Or, le produit scalaire est

$$AA'=|A|\,|A'|\cos\varphi,$$

donc

$$\cos\varphi = \frac{AA'}{|A|\,|A'|} = \frac{A_1A_1' + A_2A_2' + A_3A_3'}{\sqrt{A_1^2+A_2^2+A_3^2}\,\sqrt{A_1'^2+A_2'^2+A_3'^2}}. \tag{15}$$

Il suffit d'admettre que $0 \leqslant \varphi \leqslant \pi$.

A noter que deux plans sécants forment en fait deux dièdres $\varphi_1$ et $\varphi_2$ supplémentaires dont les cosinus sont égaux en valeur absolue mais de signe opposé ($\cos\varphi_1 = -\cos\varphi_2$). Si dans la première équation (14) on remplace $A_1$, $A_2$, $A_3$ respectivement par $-A_1$, $-A_2$, $-A_3$, on obtiendra une équation qui définit le même plan, mais dans (15) l'angle $\varphi$ est remplacé par $\pi - \varphi$.

La *condition d'orthogonalité des plans* (14) est de toute évidence $\cos\varphi = 0$, autrement dit

$$AA' = A_1A_1' + A_2A_2' + A_3A_3' = 0. \tag{16}$$

Les plans (14) sont *parallèles* si et seulement si les vecteurs $A$ et $A'$ sont colinéaires, c'est-à-dire si sont réalisées les conditions de proportionnalité

$$\frac{A_1}{A_1'} = \frac{A_2}{A_2'} = \frac{A_3}{A_3'}. \tag{17}$$

Si est remplie la condition de proportionnalité élargie

$$\frac{A_1}{A_1'} = \frac{A_2}{A_2'} = \frac{A_3}{A_3'} = \frac{B}{B'}, \tag{18}$$

alors les plans (14) sont *confondus*, c'est-à-dire que les équations (14) définissent le même plan.

Quoique la division par zéro soit impossible, il est commode de représenter les proportions (17) ou (18) avec des zéros au dénominateur. Dans ce cas, si par exemple $A_2' = 0$, on convient que $A_2 = 0$; si $B' = 0$, que $B = 0$.

Exemple 1. Les équations

$$2x + 4y + 1 = 0,$$
$$x + 2y + 5 = 0$$

définissent un couple de plans parallèles, les équations

$$2x + 4y + 2 = 0,$$
$$x + 2y + 1 = 0,$$

un couple de plans confondus.

D i s t a n c e d'u n p o i n t à u n p l a n. On demande la distance d'un point $x^0 = (x_1^0, x_2^0, x_3^0)$ à un plan $L$ défini par l'équation

$$A_1x_1 + A_2x_2 + A_3x_3 + B = 0.$$

Ramenons l'équation de $L$ à la forme normale:

$$(x, \nu) = p \qquad (p \geqslant 0).$$

Sur la figure 131 on voit que la différence $x - x^0$ entre le rayon vecteur du point courant $x$ de $L$ et le rayon vecteur de $x^0$ est un vecteur dont la valeur absolue de la projection sur $\mathbf{v}$ est égale à la distance cherchée $d$ de $x^0$ à $L$:

$$d = |(x - x^0, \mathbf{v})|.$$

Or,

$$(x - x^0, \mathbf{v}) = (x, \mathbf{v}) - (x^0, \mathbf{v}) = p - (x^0, \mathbf{v}),$$

donc

$$d = |(x^0, \mathbf{v}) - p|.$$

On voit que pour calculer la distance $d$ du point $x^0$ au plan $L$ il faut écrire l'équation de $L$ sous la forme normale, faire passer $p$ au pre-

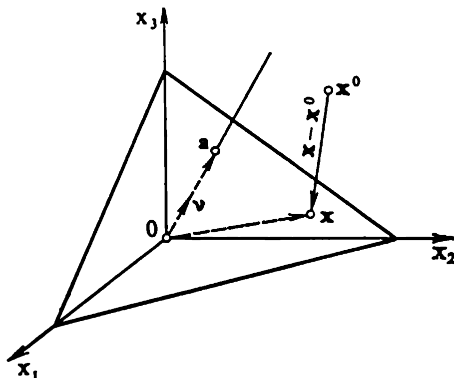



Fig. 131

mier membre et remplacer $x$ par $x^0$. Le module de l'expression obtenue est la distance cherchée $d$.

Dans le langage des paramètres du plan

$$d = \frac{|A_1x_1^0 + A_2x_2^0 + A_3x_3^0 + B|}{\sqrt{A_1^2 + A_2^2 + A_3^2}}.$$

Il est aisé de voir que si le point $x^0$ et l'origine des coordonnées sont situés de part et d'autre du plan $L$ (voir fig. 131), alors le vecteur $x - x^0$ forme avec $\mathbf{v}$ un angle obtus et par suite

$$d = -(x - x^0, \mathbf{v}) = (x^0, \mathbf{v}) - p > 0.$$

Si le point $x^0$ et l'origine des coordonnées sont situés d'un même côté de $L$ l'angle en question est aigu et

$$d = (x - x^0, \mathbf{v}) = p - (x^0, \mathbf{v}) > 0.$$

Donc, dans le premier cas $(x^0, \mathbf{v}) > p$ et dans le second $(x^0, \mathbf{v}) < p$.

EXEMPLE 2. La distance $d$ du point (1, 1, 1) au plan $L$

$$x + 2y + z = 3$$

est égale à

$$d = \left| \frac{x+2y+z-3}{\sqrt{1+4+1}} \right|_{x=1,\ y=1,\ z=1} = \frac{1}{\sqrt{6}}.$$

Le point (1, 1, 1) et l'origine des coordonnées sont situés de part et d'autre du plan $L$, puisque $M > 0$ et

$$[x + 2y + z - 3]_{x=1\ y=1\ z=1} = 1 > 0.$$

EXERCICES

1. Réduire les équations des plans

$$x-y+z=2, \quad 2x-y+\sqrt{20}\,z=10$$

à la forme normale.

2. Trouver l'angle des plans

$$x+y+z=1 \quad \text{et} \quad y=0$$
$$x-y+z=2 \quad \text{et} \quad x+y+z=3.$$

3. Former l'équation du plan passant par les points (0, 0, 1), (1, 0, 1), (1, 1, 0).

4. Former l'équation de la surface sphérique centrée en l'origine des coordonnées et tangente au plan $2x - 6y + 7z = 2$.

## § 10. La droite dans l'espace

Soit une droite $L$ dans l'espace. Repérons un point $x^0$ sur $L$ et considérons un vecteur $a = (a_1, a_2, a_3) \neq 0$ d'origine $x^0$ et de direction $L$ (fig. 132). Soit $x$ le point courant de $L$. Le vecteur $x - x^0$ peut se mettre sous la forme $x - x^0 = ta$, où $t$ est un nombre (scalaire). Si la variable réelle $t$ parcourt l'intervalle $]-\infty, \infty[$, alors $x = x^0 + ta$ parcourt la droite $L$ tout entière. On dit que

$$x - x^0 = ta, \quad t \in ]-\infty, \infty[, \qquad (1)$$

est l'*équation vectorielle d'une droite passant par* $x^0$ *et orientée dans le sens du vecteur* $a$.

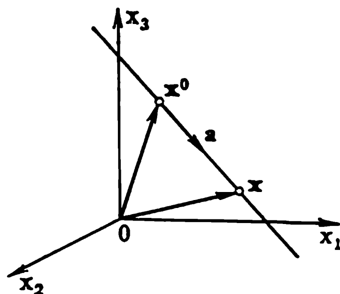


Fig. 132

Dans le langage des coordonnées, l'équation (1) se décompose en trois équations:

$$\begin{cases} x_1 - x_1^0 = ta_1, \\ x_2 - x_2^0 = ta_2, \\ x_3 - x_3^0 = ta_3. \end{cases} \qquad (1')$$

L'élimination du paramètre $t$ entre ces équations nous donne les équations de la droite (un système de deux équations)

$$\frac{x_1-x_1^0}{a_1}=\frac{x_2-x_2^0}{a_2}=\frac{x_3-x_3^0}{a_3}, \tag{1''}$$

où les nombres $a_1$, $a_2$, $a_3$ ne sont pas simultanément nuls. Les équations (1") s'appellent *équations de la droite sous la forme canonique.*

Remarque. Il est possible qu'un ou deux nombres parmi $a_1$, $a_2$, $a_3$ soient nuls. Pour la commodité on écrira les égalités (1") avec un ou deux zéros au dénominateur.

Exemple 1. Les équations

$$\frac{x-1}{1}=\frac{y-2}{0}=\frac{z-3}{2} \tag{2}$$

définissent une droite dans l'espace, passant par le point (1, 2, 3) et orientée dans le sens du vecteur (1, 0, 2).

On peut remplacer ces équations par les équations équivalentes

$$y-2=0\cdot(x-1), \qquad 2(x-1)=z-3,$$

c'est-à-dire que

$$y=2, \qquad z=2x+1. \tag{2'}$$

Donc, la droite en question est l'intersection des deux plans définis par les équations (2').

Exemple 2. Les équations de la droite

$$\frac{x-1}{1}=\frac{y-2}{0}=\frac{z-3}{0}$$

sont équivalentes aux équations

$$y-2=0, \qquad z-3=0.$$

Soient données les équations des plans

$$A_1x+A_2y+A_3z+B=0, \tag{3}$$
$$A_1'x+A_2'y+A_3'z+B'=0. \tag{4}$$

Si les coefficients $A_j$ sont proportionnels respectivement aux coefficients $A_j'$ ($A_1 : A_2 : A_3 = A_1' : A_2' : A_3'$), alors les plans (3) et (4) sont parallèles (ou confondus si $A_1 : A_2 : A_3 : B = A_1' : A_2' : A_3' : B'$) (voir § 9 (17) et (18)). Le cas échéant, les plans (3) et (4) se coupent suivant une droite. Alors l'un des déterminants

$$\begin{vmatrix} A_1 & A_2 \\ A_1' & A_2' \end{vmatrix}, \quad \begin{vmatrix} A_1 & A_3 \\ A_1' & A_3' \end{vmatrix}, \quad \begin{vmatrix} A_2 & A_3 \\ A_2' & A' \end{vmatrix}$$

est différent de zéro. Supposons pour fixer les idées qu'il s'agit du premier:

$$\begin{vmatrix} A_1 & A_2 \\ A_1' & A_2' \end{vmatrix} \neq 0.$$

Les équations (3) et (4) sont alors solubles par rapport à $x$ et $y$ et l'on obtient

$$\left\{\begin{array}{l} x = \alpha z + \mu, \\ y = \beta z + \nu, \end{array}\right. \tag{5}$$

où $\alpha$, $\beta$, $\mu$ et $\nu$ sont des nombres. Les équations (5) sont équivalentes aux suivantes:

$$\frac{x-\mu}{\alpha} = \frac{y-\nu}{\beta} = \frac{z}{1}. \tag{6}$$

On voit que la droite définie par les équations (3) et (4) passe par le point $(\mu, \nu, 0)$ et est orientée dans le sens du vecteur $(\alpha, \beta, 1)$. Si $\alpha$ ou $\beta$ ou les deux à la fois sont nuls, les équations (6) sont purement symboliques.

EXEMPLE 3. La droite définie par les équations $x = 0$, $y = 0$ est de toute évidence l'axe $Oz$. On aurait pu aboutir formellement à ce résultat. On a

$$x = 0 \cdot z, \qquad y = 0 \cdot z,$$

d'où

$$\frac{x}{0} = \frac{y}{0} = \frac{z}{1},$$

c'est-à-dire l'équation d'une droite passant par l'origine des coordonnées (0, 0, 0) dans le sens du vecteur (0, 0, 1). Il est clair que c'est l'axe $Oz$.

EXEMPLE 4. Trouver l'angle des droites

$$\frac{x_1 - x_1^1}{a_1} = \frac{x_2 - x_2^1}{a_2} = \frac{x_3 - x_3^1}{a_3}, \tag{7}$$

$$\frac{x_1 - x_1^2}{b_1} = \frac{x_2 - x_2^2}{b_2} = \frac{x_3 - x_3^2}{b_3}. \tag{8}$$

Les vecteurs $a = (a_1, a_2, a_3)$ et $b = (b_1, b_2, b_3)$ sont portés par les droites et par convention ont pour origine respectivement $x^1 = (x_1^1, x_2^1, x_3^1)$ et $x^2 = (x_1^2, x_2^2, x_3^2)$. Donc, l'angle $\varphi$ de ces vecteurs sera l'angle des droites (7) et (8):

$$\cos\varphi = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{|a|\,|b|}.$$

EXERCICES

1. Former l'équation de la droite qui passe par le point (2, —1, 0) perpendiculairement au plan $2x + z - 4y = 7$.

2. Former l'équation du plan passant par le point (1, —1, 2) et qui est orthogonal à la droite définie par les équations $2x + 3y = 1$, $3x - z = 1$.

## § 11. Orientation des systèmes de coordonnées rectangulaires

Les figures 133 et 134 représentent des systèmes de coordonnées $(Ox_1, Ox_2)$ d'*orientation contraire*. Dans la figure 133 on ne peut faire coïncider les axes $Ox_1$ et $Ox_2$ dans une rotation $(0, \pi/2)$ qu'en

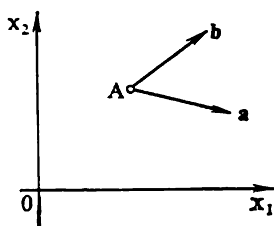



Fig. 133

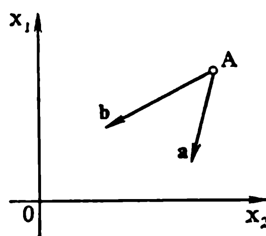



Fig. 134

faisant tourner $Ox_1$ dans le sens direct (sens qui est l'inverse de celui des aiguilles d'une montre). Sur la figure 134 cette coïncidence est réalisée par une rotation $(0, \pi/2)$ de $Ox_1$ dans le sens rétrograde. Le système de coordonnées de la figure 133 ne peut être amené en coïncidence avec celui de la figure 134 de telle manière que les axes correspondants aient le même sens.

Les figures 133 et 134 représentent un couple de vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ non colinéaires d'origine commune $A$. En déplaçant ce couple rigidement, on peut amener le point $A$ en coïncidence avec l'origine des coordonnées $O$. On se propose par une rotation de centre $O$ de faire coïncider le vecteur $\boldsymbol{a}$ avec l'axe $Ox_1$, et le vecteur $\boldsymbol{b}$ avec l'axe $Ox_2$. Ceci étant, on exige que les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ restent dans le plan considéré et que leur angle ne soit égal ni à 0 ni à $\pi$. Il est évident qu'on peut toujours réaliser cet objectif (voir fig. 135). On tourne d'abord le système rigide de vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ autour du point $O$ jusqu'à faire coïncider le vecteur $\boldsymbol{a}$ avec l'axe $Ox_1$. Les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ n'étant pas colinéaires, le vecteur $\boldsymbol{b}$ se retrouve dans le demi-plan supérieur ou inférieur. On fait pivoter le vecteur $\boldsymbol{b}$ pour l'amener en coïncidence avec l'axe $Ox_2$. Si le vecteur $\boldsymbol{b}$ est de même sens que l'axe $Ox_2$, on dit que *le couple de vecteurs* $(\boldsymbol{a}, \boldsymbol{b})$ *est de même orientation que le système de coordonnées* $(Ox_1, Ox_2)$, s'il est de sens contraire, on dit alors que *le couple* $(\boldsymbol{a}, \boldsymbol{b})$ *est d'orientation contraire à celle de* $(Ox_1, Ox_2)$.

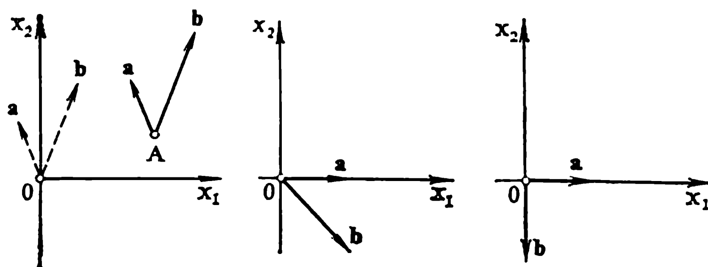



Fig. 135

Les systèmes de coordonnées rectangulaires $(Ox_1, Ox_2, Ox_3)$ des figures 136 et 137 sont différents aussi. En traitant le système de coordonnées de la figure 136 comme un système rigide on peut par un déplacement adéquat faire coïncider les axes $Ox_1$ et $Ox_2$ respectivement. Mais le sens de l'axe $Ox_3$ du premier système n'est pas le même que celui de l'axe $Ox_3$ du second. On dit que les systèmes des figures 136 et 137 sont d'*orientation contraire*. Le système

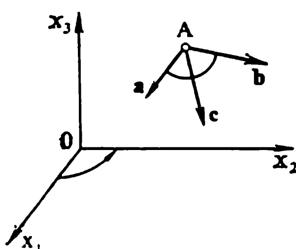



Fig. 136

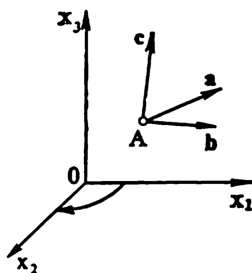



Fig. 137

de la figure 136 est *direct*, celui de la figure 28, *rétrograde*. Pour reconnaître l'orientation d'un système on peut se servir de l'un des deux procédés suivants. Un système est direct (resp. rétrograde) si une vis à filet droit (resp. gauche) tournée verticalement dans le sens de la flèche de la figure 136 (resp. fig. 137) s'enfoncera dans le sens de $O$ vers $x_3$. Un système est direct (resp. rétrograde) si la rotation d'un angle inférieur à $\pi$ autour de $Ox_3$, qui amène $Ox_1$ sur $Ox_2$, s'effectue vers la gauche (resp. droite) pour un observateur placé suivant $Ox_3$ dans le sens de $O$ vers $x_3$.

On dit que des vecteurs $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ sont *coplanaires* s'ils sont situés dans un même plan ou dans des plans parallèles.

Soit un système de vecteurs non coplanaires $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ d'origine quelconque $A$. Dans le plan engendré par les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$, faisons pivoter le vecteur $\boldsymbol{b}$ autour de $A$ jusqu'à ce qu'il soit orthogonal à $\boldsymbol{a}$. Au cours de cette rotation on prendra garde à ce que l'angle de $\boldsymbol{a}$

et de $\boldsymbol{b}$ ne soit égal ni à 0 ni à $\pi$. Ensuite faisons pivoter le vecteur $\boldsymbol{c}$ autour de $A$ pour le rendre orthogonal aux vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$. On veillera à ce que le vecteur $\boldsymbol{c}$ ne soit à aucun moment contenu dans le plan engendré par les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$. Transportons le triplet de vecteurs orthogonaux $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ en $O$ et faisons-le pivoter autour de $O$ de manière à faire coïncider les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ avec les axes $Ox_1$ et $Ox_2$. Deux cas sont possibles: 1) le vecteur $\boldsymbol{c}$ est de même sens que l'axe $Ox_3$; 2) le vecteur $\boldsymbol{c}$ est de sens contraire à celui de $Ox_3$. Dans le premier cas on dira que le système de vecteurs $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ est *de même orientation que le système de coordonnées* $(Ox_1, Ox_2, Ox_3)$, dans le second, qu'il est *d'orientation contraire* (voir fig. 137 et 136 respectivement).

## § 12. Produit vectoriel

Soient

$$\boldsymbol{a} = (a_1, a_2, a_3), \qquad \boldsymbol{b} = (b_1, b_2, b_3)$$

des vecteurs dans un système de coordonnées rectangulaires. On appelle *produit vectoriel* de $\boldsymbol{a}$ par $\boldsymbol{b}$ le vecteur $\boldsymbol{c}$:

$$\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} = [\boldsymbol{a} \times \boldsymbol{b}] = (a_2b_3 - a_3b_2)\,\boldsymbol{i} + (a_3b_1 - a_1b_3)\,\boldsymbol{j} + (a_1b_2 - a_2b_1)\,\boldsymbol{k} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \qquad (1)$$

Le dernier membre de (1) est un « déterminant généralisé » dont la première ligne est constituée des vecteurs unités $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$. Le second membre indique comment il faut comprendre ce déterminant généralité (le déterminant d'ordre trois a été développé suivant les éléments de la première ligne comme si $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ étaient des nombres).

Il est évident que $[\boldsymbol{a} \times (-\boldsymbol{b})] = -[\boldsymbol{a} \times \boldsymbol{b}]$.

On peut encore définir le produit vectoriel des vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ de la manière suivante:

1) si $\boldsymbol{a}$ et $\boldsymbol{b}$ sont colinéaires, leur produit vectoriel est nul;

b) si $\boldsymbol{a}$ et $\boldsymbol{b}$ ne sont pas colinéaires, leur produit vectoriel est un vecteur $\boldsymbol{c}$ perpendiculaire à $\boldsymbol{a}$ et à $\boldsymbol{b}$ et tel que le système $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ soit de même orientation que le système de coordonnées. Le module $|\boldsymbol{c}|$ du vecteur $\boldsymbol{c}$ est égal à

$$|\boldsymbol{c}| = |\boldsymbol{a}| \cdot |\boldsymbol{b}| \sin \omega \quad (0 \leqslant \omega \leqslant \pi), \qquad (2)$$

où $\omega$ est l'angle des vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$. En d'autres termes, le module $|\boldsymbol{c}|$ est égal à l'aire du parallélogramme construit sur les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ (fig. 138).

Ces deux définitions sont équivalentes.

En effet, si $c = 0$, alors de (1) il s'ensuit que les composantes des vecteurs $a$ et $b$ sont proportionnelles

$$a_1 : a_2 : a_3 = b_1 : b_2 : b_3,$$

c'est-à-dire que $a$ et $b$ sont colinéaires et par conséquent le produit vectoriel est nul d'après la deuxième définition.

Réciproquement, si $a$ et $b$ sont colinéaires, on a $a \times \beta = 0$ en vertu de la deuxième définition. Comme les composantes des vecteurs $a$ et $b$ sont proportionnelles, on a $c = 0$ en vertu de la première définition.

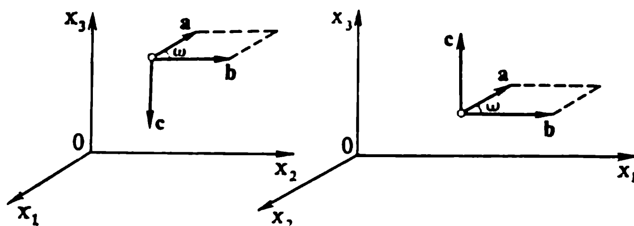



Fig. 138

Soient maintenant $a$ et $b$ des vecteurs non colinéaires et $c$ leur produit vectoriel (1). Il est évident que

$$(c, a) = (a_2b_3 - a_3b_2)\,a_1 + (a_3b_1 - a_1b_3)\,a_2 + \\ + (a_1b_2 - a_2b_1)\,a_3 = 0$$

et de façon analogue

$$(c, b) = 0.$$

Ainsi, le vecteur $c$ est orthogonal aux vecteurs $a$ et $b$.

Montrons que le système de vecteurs $(a, b, c)$ est orienté comme celui des coordonnées $(Ox_1, Ox_2, Ox_3)$. Faisons pivoter sans arrêt les vecteurs $a$ et $b$ autour du point $O$ en veillant à ce qu'ils ne soient jamais colinéaires. Donc, le vecteur $c$ ne sera jamais nul et les vecteurs $a$, $b$ et $c$, jamais coplanaires. Faisons tourner les vecteurs $a$ et $b$ de façon qu'ils aient le même sens que les axes $Ox_1$ et $Ox_2$ respectivement, c'est-à-dire qu'ils soient de la forme $a = (|a|, 0, 0)$ et $b = (0, |b|, 0)$. Ceci est possible car le plan engendré par les vecteurs $a$ et $b$ peut pivoter dans l'espace. Le vecteur $c$ calculé avec la formule (1) est alors de la forme $c = (0, 0, |a|\,|b|)$. On voit donc qu'au terme de ce processus les vecteurs $a$, $b$, $c$ sont orientés comme les axes $Ox_1$, $Ox_2$, $Ox_3$. Donc, par définition de l'orientation (voir § 11) le système initial $(a, b, c)$ est orienté comme le système de coordonnées $(Ox_1, Ox_2, Ox_3)$.

Ainsi, le produit vectoriel $c = a \times b$ défini par la formule (1) est un vecteur orthogonal aux vecteurs $a$ et $b$, et le système de vecteurs $(a, b, c)$ est orienté comme le système de coordonnées $(Ox_1, Ox_2, Ox_3)$.

Il reste à prouver la formule (2). Supposons que les vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ forment avec les axes de coordonnées des angles égaux respectivement à $\alpha_1, \alpha_2, \alpha_3$ et $\beta_1, \beta_2, \beta_3$. Comme

$$a_j = |\boldsymbol{a}| \cos \alpha_j, \quad b_j = |\boldsymbol{b}| \cos \beta_j \quad (j = 1, 2, 3),$$

il vient

$$\begin{aligned} |\boldsymbol{c}|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 = \\ = |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 [(\cos \alpha_2 \cos \beta_3 - \cos \alpha_3 \cos \beta_2)^2 + \\ + (\cos \alpha_3 \cos \beta_1 - \cos \alpha_1 \cos \beta_3)^2 + \\ + (\cos \alpha_1 \cos \beta_2 - \cos \alpha_2 \cos \beta_1)^2] = \\ = |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 [(\cos^2 \alpha_1 + \cos^2 \alpha_2 + \cos^2 \alpha_3) \times \\ \times (\cos^2 \beta_1 + \cos^2 \beta_2 + \cos^2 \beta_3) - \\ - (\cos \alpha_1 \cos \beta_1 + \cos \alpha_2 \cos \beta_2 + \cos \alpha_3 \cos \beta_3)^2] = \\ = |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 [1 \cdot 1 - \cos^2 \omega] = |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 \sin^2 \omega, \end{aligned}$$

où $\omega$ est l'angle des vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$

$$\left(\cos \omega = \sum_{j=1}^{3} \cos \alpha_j \cos \beta_j\right).$$

Ceci prouve (2).

Donc, nous avons entièrement démontré que la définition du produit vectoriel à l'aide de la formule (1) entraîne la deuxième définition.

Réciproquement, si un vecteur est justiciable de la deuxième définition, il est unique, car il n'existe qu'un seul vecteur orthogonal à $\boldsymbol{a}$ et à $\boldsymbol{b}$, dont la longueur soit égale à l'aire du parallélogramme construit sur $\boldsymbol{a}$ et $\boldsymbol{b}$, et tel que le système $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ soit orienté comme le système $(Ox_1, Ox_2, Ox_3)$. Donc, ce vecteur n'est autre que le vecteur $\boldsymbol{c}$ défini par la formule (1).

Soulignons une fois de plus que la condition $\boldsymbol{a} \times \boldsymbol{b} = 0$ est une *condition nécessaire et suffisante de colinéarité des vecteurs* $\boldsymbol{a}$ *et* $\boldsymbol{b}$.

Considérons maintenant deux vecteurs non nuls

$$\boldsymbol{a} = (a_1, a_2), \quad \boldsymbol{b} = (b_1, b_2) \tag{3}$$

dans un système de coordonnées rectangulaires $(Ox_1, Ox_2)$ (fig. 139, 140). Traçons un axe $Ox_3$ orthogonal aux axes $Ox_1$ et $Ox_2$. Les coordonnées des vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ seront maintenant

$$\boldsymbol{a} = (a_1, a_2, 0), \quad \boldsymbol{b} = (b_1, b_2, 0).$$

Le produit vectoriel de ces vecteurs est égal à

$$c = \boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \boldsymbol{k}, \tag{4}$$

où $k$ est le vecteur unité de l'axe $Ox_3$. Par définition du produit vectoriel le système ($a$, $b$, $c$) est orienté comme le système de coordonnées ($Ox_1$, $Ox_2$, $Ox_3$). Donc, il est évident que le système ($a$, $b$) est de même orientation ou d'orientation contraire à celle du système de coordonnées ($Ox_1$, $Ox_2$) selon que le déterminant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

est positif ou négatif. Le premier cas est représenté sur la figure 139, le second, sur la figure 140. On sait également que l'aire du parallé-

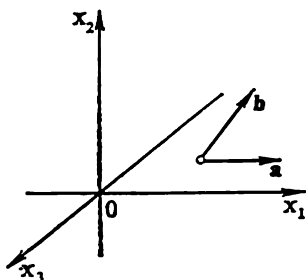



Fig. 139

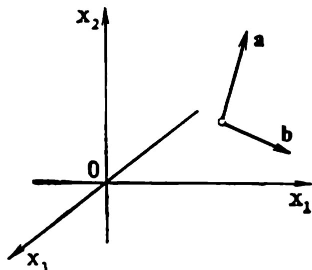



Fig. 140

logramme construit sur les vecteurs $a$ et $b$ est égale à (voir (4))

$$S = |c| = \left\| \begin{matrix} a_1 & a_2 \\ b_1 & b_2 \end{matrix} \right\|,$$

c'est-à-dire à la valeur absolue du déterminant

$$\Delta = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}. \tag{5}$$

Donc, nous avons prouvé que: 1) la valeur absolue du déterminant (5) est égale à l'aire du parallélogramme construit sur les vecteurs $a = (a_1, a_2)$ et $b = (b_1, b_2)$ ; 2) le signe du déterminant (5) dépend de la position de ces vecteurs par rapport au système de coordonnées ($Ox_1$, $Ox_2$). On a le signe + si le système ($a$, $b$) est orienté comme ($Ox_1$, $Ox_2$) et le signe — dans le cas contraire.

On a les relations

$$a \times b = -[b \times a], \tag{6}$$

$$a \times (\alpha b) = \alpha [a \times b], \tag{7}$$

$$a \times (b + c) = [a \times b] + [a \times c], \tag{8}$$

où $a$, $b$, $c$ sont des vecteurs quelconques, $\alpha$ un scalaire.

On établit facilement ces relations en exprimant les produits vectoriels par la formule (1) en fonction des coordonnées des vecteurs

$$a = (a_1, a_2, a_3), \quad b = (b_1, b_2, b_3), \quad c = (c_1, c_2, c_3).$$

Les formules (6) et (7) peuvent être également établies sans peine par des considérations géométriques. Soient $\boldsymbol{a}$ et $\boldsymbol{b}$ des vecteurs non colinéaires. Si dans le produit vectoriel on intervertit $\boldsymbol{a}$ et $\boldsymbol{b}$, l'aire du parallélogramme construit sur $\boldsymbol{a}$ et $\boldsymbol{b}$ et la perpendiculaire à $\boldsymbol{a}$ et $\boldsymbol{b}$ ne changeront pas. Seul changera le sens de $\boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b}$ et par suite l'orientation de $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} \times \boldsymbol{b})$.

La multiplication du vecteur $\boldsymbol{b}$ par un nombre positif $\alpha$ augmente de $\alpha$ fois l'aire du parallélogramme construit sur $\boldsymbol{a}$ et $\boldsymbol{b}$ et ne change pas le sens du produit vectoriel. Si $\alpha < 0$, alors

$$\boldsymbol{a} \times (\alpha \boldsymbol{b}) = \boldsymbol{a} \times (-|\alpha| \boldsymbol{b}) = |\alpha| [\boldsymbol{a} \times (-\boldsymbol{b})] = \\ = -|\alpha| [\boldsymbol{a} \times \boldsymbol{b}] = \alpha [\boldsymbol{a} \times \boldsymbol{b}].$$

On remarquera en outre que de (6) et de (7) il résulte que

$$(\alpha \boldsymbol{a}) \times \boldsymbol{b} = -[\boldsymbol{b} \times (\alpha \boldsymbol{a})] = -\alpha [\boldsymbol{b} \times \boldsymbol{a}] = \alpha [\boldsymbol{a} \times \boldsymbol{b}].$$

Exemple. Trouver l'angle $A$ du triangle $ABC$ de sommets $A = (1, 2, 3)$, $B = (2, 2, 2)$, $C = (1, 2, 4)$.

Désignons l'angle cherché par $\omega$. L'angle $\omega$ est celui des vecteurs $\overrightarrow{AB}$ et $\overrightarrow{AC}$. La deuxième définition du produit vectoriel donne

$$\sin \omega = \frac{|\overrightarrow{AB} \times \overrightarrow{AC}|}{|\overrightarrow{AB}|\,|\overrightarrow{AC}|},$$

où $\overrightarrow{AB} = (1, 0, -1)$, $\overrightarrow{AC} = (0, 0, 1)$; $|\overrightarrow{AB}| = \sqrt{2}$, $|\overrightarrow{AC}| = 1$,

$$\overrightarrow{AB} \times \overrightarrow{AC} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{vmatrix} = -\boldsymbol{j}.$$

D'où

$$\sin \omega = \frac{1}{\sqrt{2}}, \quad \omega = \frac{\pi}{4}, \quad \omega = \frac{3\pi}{4}.$$

Comme $\overrightarrow{BC} = (-1, 0, 2)$ et $|\overrightarrow{BC}|^2 = 5 > |\overrightarrow{AC}|^2 + |\overrightarrow{AB}|^2 = 1 + 2 = 3$, on a $\omega = 3\pi/4$.

Remarque. Si l'angle $A$ du triangle $ABC$ est droit, alors $|\overrightarrow{BC}|^2 = BC^2 = AC^2 + AB^2$; si $A$ est obtus, alors $BC^2 > AC^2 + AB^2$; si $A$ est aigu, alors $BC^2 < AC^2 + AB^2$.

## § 13. Produit mixte

On appelle *produit mixte des vecteurs* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ le nombre égal au produit scalaire du vecteur $\boldsymbol{a} \times \boldsymbol{b}$ par le vecteur $\boldsymbol{c}$:

$$(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c} = (a_2b_3 - a_3b_2)\, c_1 + (a_3b_1 - a_1b_3)\, c_2 + (a_1b_2 - a_2b_1)\, c_3 = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \quad (1)$$

Par définition du produit scalaire on a

$$(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c} = |\, \boldsymbol{a} \times \boldsymbol{b} \,|\, \mathrm{pr}_{\boldsymbol{a}\times\boldsymbol{b}}\boldsymbol{c} = (|\, \boldsymbol{a} \,|\cdot|\, \boldsymbol{b} \,| \sin \omega)\, \mathrm{pr}_{\boldsymbol{a}\times\boldsymbol{b}}\boldsymbol{c}.$$

Donc, on peut de toute évidence dire que le produit mixte $(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c}$ est égal au volume du parallélépipède construit sur les vecteurs $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, pris avec le signe $+$ ou $-$ selon que le système $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ est de même orientation ou d'orientation contraire à $(Ox_1, Ox_2, Ox_3)$. A noter que $|\, \mathrm{pr}_{\boldsymbol{a}\times\boldsymbol{b}}\boldsymbol{c} \,|$ est égale à la hauteur du parallélépipède.

On a les relations

$$(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c} = (\boldsymbol{c} \times \boldsymbol{a})\, \boldsymbol{b} = (\boldsymbol{b} \times \boldsymbol{c})\, \boldsymbol{a} \quad (2)$$

qui sont une conséquence immédiate des propriétés du déterminant (1).

Si les vecteurs $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ sont coplanaires, alors

$$(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c} = 0,$$

puisque $\boldsymbol{a} \times \boldsymbol{b}$ est orthogonal à $\boldsymbol{c}$. Inversement, si $(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c} = 0$, alors le vecteur $\boldsymbol{c}$ est orthogonal à $\boldsymbol{a} \times \boldsymbol{b}$ et par suite est situé dans le plan des vecteurs $\boldsymbol{a}$ et $\boldsymbol{b}$ ou dans un plan parallèle.

Donc, la condition

$$(\boldsymbol{a} \times \boldsymbol{b})\, \boldsymbol{c} = 0$$

est une *condition nécessaire et suffisante de coplanation des vecteurs* $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$.

EXEMPLE. Etablir la condition de coplanation de quatre points.

Soient donnés quatre points $A_j = (x_j, y_j, z_j)$ $(j = 1, 2, 3, 4)$. Si ces points sont coplanaires, les vecteurs $\overrightarrow{A_1A_2}$, $\overrightarrow{A_1A_3}$, $\overrightarrow{A_1A_4}$ le seront aussi et par suite leur produit mixte est nul, soit

$$(\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3})\, \overrightarrow{A_1A_4} = \begin{vmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ x_4 - x_1 & y_4 - y_1 & z_4 - z_1 \end{vmatrix} = 0.$$

Ceci est la condition de coplanation de quatre points (cf. § 9, (12)).

## § 14. Système de vecteurs linéairement indépendants

Soit dans $R_n$ un système de $k$ vecteurs

$$a^s = (a_{s1}, a_{s2}, \ldots, a_{sn}) \quad (s = 1, \ldots, k). \tag{1}$$

Par définition le système (1) est *linéairement indépendant* si l'égalité vectorielle

$$\lambda_1 a^1 + \lambda_2 a^2 + \ldots + \lambda_k a^k = 0, \tag{2}$$

où $\lambda_1, \lambda_2, \ldots, \lambda_k$ sont des nombres, entraîne

$$\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

On dit que le système de vecteurs (1) est *linéairement dépendant* s'il existe des nombres $\lambda_1, \lambda_2, \ldots, \lambda_k$ non tous nuls vérifiant l'égalité (2). Si, pour fixer les idées, l'on admet que $\lambda_k \neq 0$, il s'ensuit de (2) que

$$a^k = \mu_1 a^1 + \ldots + \mu_{k-1} a^{k-1}, \tag{3}$$

où

$$\mu_1 = -\frac{\lambda_1}{\lambda_k}, \ \ldots, \ \mu_{k-1} = -\frac{\lambda_{k-1}}{\lambda_k}.$$

*Si donc un système de $k$ vecteurs est linéairement dépendant, l'un de ces vecteurs est une combinaison linéaire des autres.*

Comme il sera toujours question dans la suite de dépendance *linéaire*, on omettra parfois le terme *linéaire*. On dira aussi *vecteurs dépendants* et *indépendants* au lieu de *système de vecteurs dépendants* et *indépendants*.

Un vecteur $a^1$ forme un système linéairement indépendant s'il est différent de 0 et linéairement dépendant s'il est égal à 0.

Si un système de vecteurs $a^1, \ldots, a^k$ est linéairement indépendant, tout sous-système $a^1, \ldots, a^s$ l'est à fortiori ($s < k$). Dans le cas contraire il existerait un système de nombres $\lambda_1, \ldots, \lambda_s$ pour lequel

$$\lambda_1 a^1 + \ldots + \lambda_s a^s = 0,$$

et pour le système non trivial $\lambda_1, \ldots, \lambda_s, \overbrace{0, \ldots, 0}^{k-s \text{ fois}}$, on aurait

$$\lambda_1 a^1 + \ldots + \lambda_s a^s + 0 \cdot a^{s+1} + \ldots + 0 \cdot a^k = 0.$$

De ce qui précède il s'ensuit que si le système de vecteurs $a^1, \ldots, a^s$ est linéairement dépendant, tout sur-système

$$a^1, \ldots, a^s, a^{s+1}, \ldots, a^k$$

l'est également. En particulier, un système de vecteurs contenant le vecteur 0 est toujours linéairement dépendant.

Formons la matrice du système (1)

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{Vmatrix}.$$

THÉORÈME. *Le système* (1) *est linéairement indépendant si* rang $A = k$, *c'est-à-dire si* rang $A$ *est égal au nombre de vecteurs.*
*Si* rang $A < k$, *le système* (1) *est linéairement dépendant.*

EXEMPLE 1. Deux vecteurs $\boldsymbol{a}^1 = (a_{11}, a_{12})$, $\boldsymbol{a}^2 = (a_{21}, a_{22})$ de $R_2$ sont linéairement indépendants si le déterminant

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0,$$

car l'équation vectorielle

$$\lambda_1 \boldsymbol{a}^1 + \lambda_2 \boldsymbol{a}^2 = 0 \tag{4}$$

est équivalente aux deux équations en les coordonnées

$$\left\{ \begin{array}{l} a_{11}\lambda_1 + a_{21}\lambda_2 = 0, \\ a_{12}\lambda_1 + a_{22}\lambda_1 = 0. \end{array} \right. \tag{5}$$

Or, si $\Delta \neq 0$, le système (5) admet une solution unique trivial

$$\lambda_1 = \lambda_2 = 0. \tag{6}$$

Si $\Delta = 0$, les équations (5) sont satisfaites par un système non trivial $(\lambda_1, \lambda_2)$, c'est-à-dire que pour $\Delta = 0$ les vecteurs $\boldsymbol{a}^1$, $\boldsymbol{a}^2$ sont linéairement dépendants.

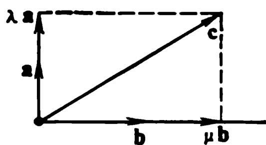


Fig. 141

Il est évident que dans $R_2$ il y a équivalence entre la colinéarité et la dépendance linéaire des vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$, de même qu'entre la non-colinéarité et l'indépendance linéaire.

EXEMPLE. 2. Dans $R_2$ tout système de vecteurs $\boldsymbol{a}^1, \boldsymbol{a}^2, \ldots, \boldsymbol{a}^k$ $(k \geqslant 3)$ est linéairement dépendant. Géométriquement, cela ressort de la figure 141 : si $\boldsymbol{c}$ est un vecteur arbitraire et $\boldsymbol{a}$ et $\boldsymbol{b}$ sont des vecteurs non colinéaires, on peut toujours exhiber des nombres $\lambda$ et $\mu$ tels que

$$\boldsymbol{c} = \lambda \boldsymbol{a} + \mu \boldsymbol{b}.$$

Ceci montre que les vecteurs $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ sont linéairement dépendants. Si $\boldsymbol{a}$ et $\boldsymbol{b}$ sont des vecteurs colinéaires, ils sont linéairement dépendants et à plus forte raison $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ le seront.

D'après le théorème ci-dessus, pour étudier un couple de vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$, il faut composer la matrice de leurs coordonnées, soit

$$A = \left\| \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix} \right\|.$$

Dans ce cas $k = 2$.

a) Si rang $A = 1 < 2 = k$, le théorème nous dit que les vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$ sont linéairement dépendants.

b) Si rang $A = 2 = k$, les vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$ sont linéairement indépendants.

Ceci confirme les résultats obtenus, car dans le cas a) $\Delta = 0$ et dans b) $\Delta \neq 0$.

Le fait que trois vecteurs quelconques $a^1$, $a^2$, $a^3$ de $R_2$ sont linéairement dépendants est également prévu par le théorème. En effet,

$$\operatorname{rang} A \leqslant 2 < 3 = k.$$

Exemple 3. Dans l'espace $R_3$ les vecteurs

$$a^1 = (a_{11}, \ a_{12}, \ a_{13}), \quad a^2 = (a_{21}, \ a_{22}, \ a_{23})$$

sont linéairement dépendants si et seulement s'ils sont colinéaires.

En effet, supposons que $a^1$ et $a^2$ sont colinéaires. Si l'un de ces vecteurs est nul, ils sont linéairement dépendants. Si $a^1$ et $a^2$ sont colinéaires et non nuls, alors

$$a^1 = \lambda a^2,$$

où $\lambda$ est un nombre. Ceci exprime que $a^1$ et $a^2$ sont linéairement dépendants.

Réciproquement, si $a^1$ et $a^2$ sont linéairement dépendants, l'un d'eux dépend de l'autre, par exemple

$$a^2 = \mu a^1,$$

c'est-à-dire que ces vecteurs sont colinéaires.

Si l'on considère la matrice

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{Vmatrix},$$

on constate que les éléments des lignes sont proportionnels, donc

$$\operatorname{rang} A = 1 < 2 = k,$$

c'est-à-dire que notre proposition concorde avec le théorème 1.

Exemple 4. Considérons maintenant trois vecteurs dans $R_3$:

$$a^1 = (a_{11}, \ a_{12}, \ a_{13}),$$
$$a^2 = (a_{21}, \ a_{22}, \ a_{23}),$$
$$a^3 = (a_{31}, \ a_{32}, \ a_{33}).$$

Soit

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

L'équation vectorielle

$$\lambda_1 a^1 + \lambda_2 a^2 + \lambda_3 a^3 = 0 \tag{7}$$

est équivalente au système

$$\begin{cases} a_{11}\lambda_1 + a_{21}\lambda_2 + a_{31}\lambda_3 = 0, \\ a_{12}\lambda_1 + a_{22}\lambda_2 + a_{32}\lambda_3 = 0, \\ a_{13}\lambda_1 + a_{23}\lambda_2 + a_{33}\lambda_3 = 0. \end{cases} \tag{7'}$$

Si $\Delta \neq 0$, le système (7′) admet une solution triviale unique $\lambda_1 = \lambda_2 = \lambda_3 = 0$. L'équation (7) admet alors une solution triviale unique et les vecteurs $a^1$, $a^2$, $a^3$ sont linéairement indépendants.

Si $\Delta = 0$, le système (7') et donc l'équation (7) possèdent une solution non triviale $(\lambda_1, \lambda_2, \lambda_3)$. Par suite les vecteurs $\boldsymbol{a}^1, \boldsymbol{a}^2, \boldsymbol{a}^3$ sont linéairement dépendants. Distinguons les cas suivants:

1) rang $A = 1$, où

$$A = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}.$$

Alors l'une au moins des lignes de $A$, disons la première, possède un élément non nul. Considérons la matrice

$$A' = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{Vmatrix}. \tag{8}$$

Elle est de rang 1, donc tous ses déterminants d'ordre deux sont nuls:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = 0.$$

Il est alors évident que les composantes des vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$ sont proportionnelles

$$\frac{a_{21}}{a_{11}} = \frac{a_{22}}{a_{12}} = \frac{a_{23}}{a_{13}} = \lambda,$$

d'où

$$a_{21} = a_{11}\lambda, \quad a_{22} = a_{12}\lambda, \quad a_{23} = a_{13}\lambda$$

ou

$$\boldsymbol{a}^2 = \lambda \boldsymbol{a}^1.$$

De façon analogue, pour la matrice

$$A'' = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}$$

dont tous les déterminants d'ordre deux sont nuls, on obtient

$$\boldsymbol{a}^3 = \mu \boldsymbol{a}^1,$$

où $\mu$ est un nombre. Donc, les vecteurs $\boldsymbol{a}^1, \boldsymbol{a}^2, \boldsymbol{a}^3$ sont colinéaires.

2) rang $A = 2$. L'une, donc, des matrices formées avec deux lignes de la matrice $A$ est de rang 2. Supposons pour fixer les idées qu'il s'agit de la matrice $A'$ (voir (8)). D'après l'exemple 3, les vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$ sont linéairement indépendants. Or, le système $(\boldsymbol{a}^1, \boldsymbol{a}^2, \boldsymbol{a}^3)$ est dépendant, c'est-à-dire que pour un triplet non trivial de nombres $(\lambda_1, \lambda_2, \lambda_3)$, on a

$$\lambda_1 \boldsymbol{a}^1 + \lambda_2 \boldsymbol{a}^2 + \lambda_3 \boldsymbol{a}^3 = 0. \tag{9}$$

$\lambda_3 \neq 0$, sinon $\lambda_1 \boldsymbol{a}^1 + \lambda_2 \boldsymbol{a}^2 = 0$ et, en vertu de l'indépendance du système $\boldsymbol{a}^1, \boldsymbol{a}^2$, on aurait $\lambda_1 = \lambda_2 = 0$. Donc, l'équation (9) est soluble par rapport à $\boldsymbol{a}^3$:

$$\boldsymbol{a}^3 = \mu_1 \boldsymbol{a}^1 + \mu_2 \boldsymbol{a}^2, \quad \mu_1 = -\frac{\lambda_1}{\lambda_3}, \quad \mu_2 = -\frac{\lambda_2}{\lambda_3}.$$

Par conséquent, si $\Delta = 0$ et rang $A' = 2$ (voir (8)), alors les vecteurs $\boldsymbol{a}^1$ et $\boldsymbol{a}^2$ ne sont pas colinéaires et le vecteur $\boldsymbol{a}^3$ est situé dans le plan engendré par ces vecteurs.

Ces raisonnements ne contredisent pas le théorème. En effet, si $\Delta \neq 0$, alors rang $A = 3 = k$ et les vecteurs $\boldsymbol{a}^1$, $\boldsymbol{a}^2$, $\boldsymbol{a}^3$ sont linéairement indépendants en vortu de ce théorème. Si $\Delta = 0$, alors rang $A < 3$ et les vecteurs $\boldsymbol{a}^1$, $\boldsymbol{a}^2$, $\boldsymbol{a}^3$ sont linéairement dépendants.

Démonstration du théorème. L'égalité vectorielle (2) est équivalente aux $n$ équations suivantes en les coordonnées des vecteurs:

$$\left\{\begin{array}{l} \lambda_1 a_{11} + \lambda_2 a_{21} + \ldots + \lambda_k a_{k1} = 0, \\ \lambda_1 a_{12} + \lambda_2 a_{22} + \ldots + \lambda_k a_{k2} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1 a_{1n} + \lambda_2 a_{2n} + \ldots + \lambda_k a_{kn} = 0. \end{array}\right. \qquad (2')$$

Supposons que rang $A = k$. Il est alors évident que $k \leqslant n$. Il existe un déterminant non nul d'ordre $k$ engendré par la matrice $A$. Sans nuire à la généralité, on peut admettre que ce déterminant est formé des coefficients des $k$ premières équations du système (2'):

$$\left\{\begin{array}{l} \lambda_1 a_{11} + \ldots + \lambda_k a_{k1} = 0, \\ \ldots\ldots\ldots\ldots\ldots \\ \lambda_1 a_{1k} + \ldots + \lambda_k a_{kk} = 0, \end{array}\right. \qquad (10)$$

$$\begin{vmatrix} a_{11} & \ldots & a_{k1} \\ \ldots & \ldots & \ldots \\ a_{1k} & \ldots & a_{kk} \end{vmatrix} \neq 0.$$

Le système homogène (10), ayant un déterminant non nul, n'admet que la solution triviale

$$\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0,$$

et le système de vecteurs (1) est linéairement indépendant. Ceci achève la démonstration du théorème pour ce cas.

Supposons maintenant que rang $A = s < k$. Renumérotons les variables $\lambda_j$ de (2') de telle sorte que

$$\begin{vmatrix} a_{11} & \ldots & a_{s1} \\ \ldots & \ldots & \ldots \\ a_{1s} & \ldots & a_{ss} \end{vmatrix} \neq 0.$$

On peut alors mettre le système des $s$ premières équations sous la forme:

$$\left\{\begin{array}{l} a_{11}\lambda_1 + \ldots + a_{s1}\lambda_s = -a_{s+1,\,1}\lambda_{s+1} - \ldots - a_{k1}\lambda_k, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1s}\lambda_1 + \ldots + a_{ss}\lambda_s = -a_{s+1,\,s}\lambda_{s+1} - \ldots - a_{ks}\lambda_k. \end{array}\right.$$

Si l'on pose $\lambda_{s+1} = \ldots = \lambda_k = 1$, on peut résoudre ce système par rapport à $\lambda_1, \ldots, \lambda_s$. En définitive on obtient la solution non triviale

$$\lambda_1, \ldots, \lambda_k \tag{11}$$

des $s$ premières équations de (2'). On peut maintenant ajouter à ces $s$ équations $k - s$ autres équations quelconques du système (2'), et le système obtenu admettra la solution non triviale (11) pour solution. En effet, mettons formellement le système (2') sous la forme suivante:

$$\left\{\begin{array}{l} \lambda_1 a_{11} + \ldots + \lambda_k a_{k1} + \lambda_{k+1} \cdot 0 + \ldots + \lambda_n \cdot 0 = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \lambda_1 a_{1n} + \ldots + \lambda_k a_{kn} + \lambda_{k+1} \cdot 0 + \ldots + \lambda_n \cdot 0 = 0. \end{array}\right. \tag{2'}$$

La matrice des coefficients des équations obtenues est de rang $s$, de même d'ailleurs que la matrice élargie (c'est-à-dire la matrice formée en ajoutant à la première une colonne de zéros à droite). Les $s$ premières équations du système (2') sont satisfaites par les nombres $\lambda_1, \ldots, \lambda_k$ (voir (11)) et par des nombres arbitraires $\lambda_{k+1}, \ldots, \lambda_n$. D'après la proposition 2), § 4, les nombres $\lambda_1, \ldots$ $\ldots, \lambda_n$ vérifient les autres équations du système (2'), autrement dit les nombres $\lambda_1, \ldots, \lambda_k$ (qui ne sont pas tous nuls) satisfont aux autres équations du système (2).

Donc, les vecteurs $a^1, \ldots, a^k$ sont linéairement dépendants et le théorème est prouvé dans ce cas.

## § 15. Opérateurs linéaires

Soit une matrice carrée arbitraire

$$A = \| a_{kl} \| = \left\| \begin{array}{ccc} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\cdot\cdot\cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{array} \right\|. \tag{1}$$

On peut traiter la matrice $A$ comme un *opérateur* associant à tout vecteur $x = (x_1, \ldots, x_n) \in R_n$ le vecteur $y = (y_1, \ldots, y_n) \in R_n$ dont les coordonnées sont données par les formules

$$\left\{\begin{array}{l} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n, \\ \ldots\ldots\ldots\ldots\ldots \\ y_n = a_{n1}x_1 + \ldots + a_{nn}x_n, \end{array}\right. \tag{2}$$

ou de façon plus concise

$$y_i = \sum_{j=1}^{n} a_{ij}x_j \quad (i = 1, \ldots, n), \tag{2'}$$

c'est-à-dire que la $i$-ième coordonnée du vecteur $y$ s'exprime à l'aide de la $i$-ième ligne. Cet opérateur sera noté brièvement par:

$$y = Ax \qquad (x, y \in R_n). \tag{2''}$$

En particulier, dans le cas de l'espace à une dimension $R_1$, les vecteurs $x$ et $y$ sont des nombres et l'opérateur (2'') se transforme

en la fonction

$$y = a_{11}x.$$

*L'opérateur* (2") *est linéaire.* Cela signifie qu'il vérifie la relation

$$A(\alpha x + \beta x') = \alpha Ax + \beta Ax',$$

quels que soient les vecteurs $x$, $x' \in R_n$ et les nombres $\alpha$, $\beta$. En effet

$$\sum_{j=1}^{n} a_{ij}(\alpha x_j + \beta x'_j) = \alpha \sum_{j=1}^{n} a_{ij}x_j + \beta \sum_{j=1}^{n} a_{ij}x'_j \quad (i = 1, \ldots, n).$$

Si des matrices $A = \| a_{kl} \|$ et $B = \| b_{kl} \|$ sont égales, c'est-à-dire que leurs éléments correspondants $a_{kl}$ et $b_{kl}$ sont égaux, alors elles définissent des opérateurs identiquement égaux:

$$Ax = Bx, \qquad \forall x \in R_n. \tag{3}$$

Inversement, de l'égalité (3) il s'ensuit que

$$a_{kl} = b_{kl} \qquad (k, l = 1, \ldots, n),$$

c'est-à-dire que les matrices $A$ et $B$ sont égales. Ce qu'on vérifie immédiatement en posant dans (3)

$$x = x^l = (0, \ldots, 0, 1, 0, \ldots, 0),$$

où 1 se trouve au $l$-ième rang ($l = 1, \ldots, n$).

Donc, à des matrices différentes sont associés des opérateurs différents. Si deux matrices $A_1$ et $A_2$ ne diffèrent que par un élément au moins, il existe nécessairement un vecteur $x$ tel que

$$A_1x \neq A_2x.$$

Soit donné en plus de $A$ un opérateur $B$ défini par la matrice carrée d'ordre $n$

$$B = \| b_{kl} \|.$$

A tout $x \in R_n$ l'opérateur $A$ associe un vecteur $y \in R_n$. L'image de $y$ par l'opérateur $B$ est un vecteur $z$ dont les coordonnées sont données par les formules

$$z_k = \sum_{i=1}^{n} b_{ki}y_i \qquad (k = 1, \ldots, n).$$

On obtient en définitive l'opérateur linéaire composé

$$z = BAx \qquad (x \in R_n), \tag{4}$$

où

$$z_k = \sum_{i=1}^{n} b_{ki}y_i = \sum_{i=1}^{n} b_{ki} \sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} b_{ki}a_{ij} \right) x_j = \sum_{j=1}^{n} \gamma_{kj}x_j$$

dont la matrice $\| \gamma_{kj} \|$ s'appelle *produit des matrices B* et *A* et se note:

$$BA = \| \gamma_{kj} \|, \tag{5}$$

où

$$\gamma_{kj} = \sum_{i=1}^{n} b_{ki} a_{ij} \qquad (k, j = 1, \ldots, n), \tag{6}$$

autrement dit, l'élément $\gamma_{kj}$ de la matrice $BA$ est égal à la somme des produits des éléments de la $k$-ième ligne de la matrice $B$ par les éléments correspondants de la $j$-ième colonne de la matrice $A$.

*Le déterminant de la matrice BA est égal au produit des déterminants des matrices B et A*:

$$| BA | = | B | \, | A |. \tag{7}$$

Cette propriété est une conséquence de la formule du produit des déterminants (voir § 2, propriété $j$)).

Supposons que le déterminant de la matrice de l'opérateur $A$ (voir (1)) n'est pas nul:

$$\Delta = | a_{kl} | \neq 0.$$

Dans ce cas (voir § 4, théorème 1), le système d'équations (2), ou ce qui revient au même, l'équation opératorielle $y = Ax$ possède une solution unique $x \in R_n$ pour tout $y \in R_n$ donné. Ceci étant, les formules qui permettent de calculer $x$ pour $y \in R_n$ donné sont de la forme

$$x_j = \sum_{s=1}^{n} b_{js} y_s \qquad (j = 1, \ldots, n), \tag{8}$$

où

$$b_{js} = A_{sj}/\Delta \quad (s, j = 1, \ldots, n) \tag{9}$$

(voir § 4, (3')), $A_{sj}$ étant le cofacteur de l'élément $a_{sj}$ du déterminant $\Delta$.

Du reste, pour l'instant il importe seulement de signaler que les nombres $b_{js}$ sont les éléments de la matrice

$$B = \| b_{js} \|$$

qui est douée des remarquables propriétés suivantes:

$$BAx = x, \quad \forall x \in R_n, \tag{10}$$
$$ABy = y, \quad \forall y \in R_n. \tag{11}$$

En effet, l'image d'un vecteur $x \in R_n$ par l'opérateur $A$ est un vecteur $y$ dont l'image par $B$ est $x$. Par ailleurs, l'image de tout $y \in R_n$ par l'opérateur $B$ (voir (8)) est un vecteur $x$ tel que $Ax = y$.

Dans la relation (11), on peut de toute évidence remplacer $y$ par $x$ et obtenir l'identité

$$BAx = ABx = x, \ \forall x.$$

L'opérateur $E$ tel que $x = Ex$ ($\forall x \in R_n$) s'appelle *opérateur identique*. La matrice qui lui correspond est de la forme

$$E = \begin{Vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{Vmatrix}$$

et s'appelle *matrice unité*. Nous avons montré que

$$AB = BA = E.$$

L'opérateur $B$ vérifiant la relation précédente s'appelle *inverse* de $A$ et se note $A^{-1}$. Sa matrice se note également par $A^{-1}$. Les éléments de la matrice $A^{-1}$ s'expriment en fonction de ceux de la matrice $A$ par les formules (9).

Nous avons montré que *si le déterminant* $|A|$ *d'une matrice carrée* $A$ *n'est pas nul, alors celle-ci admet une matrice inverse* $A^{-1}$. Pour $A^{-1}$ on a donc

$$A^{-1}A = AA^{-1} = E.$$

Si le déterminant de la matrice $A$ est nul, celle-ci ne possède pas d'inverse. Ceci revient à dire que l'équation $y = Ax$ n'admet pas de solution pour tout $y$.

Cela étant, la relation $AA^{-1}y = y$, si elle a lieu, affirme que *tout* $y \in R_n$ a pour image (par l'opérateur $A^{-1}$) un $x$ qui est solution de l'équation $y = Ax$.

Remarque. La multiplication des matrices peut être étendue à des matrices non carrées $B$ et $A$, pourvu que le nombre des colonnes de $B$ soit égal à celui des lignes de $A$. La multiplication s'effectue avec des formules semblables à (6). Si par exemple

$$B = \begin{Vmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Vmatrix} \quad \text{et} \quad A = \begin{Vmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{Vmatrix},$$

alors

$$BA = \begin{Vmatrix} 4 & 3 \\ 0 & 1 \\ 1 & 0 \end{Vmatrix}.$$

Le produit $AB$ n'a pas de sens ici, car la matrice $A$ possède deux colonnes et la matrice $B$ trois lignes.

Pour les matrices carrées $A$ et $B$, les produits $AB$ et $BA$ ont un sens mais $AB$ n'est pas toujours égal à $BA$. Par exemple, pour

$$A=\begin{Vmatrix}1 & 1\\ 1 & 1\end{Vmatrix}, \quad B=\begin{Vmatrix}0 & 1\\ 0 & 0\end{Vmatrix}$$

on a

$$AB=\begin{Vmatrix}0 & 1\\ 0 & 1\end{Vmatrix}, \qquad BA=\begin{Vmatrix}1 & 1\\ 0 & 0\end{Vmatrix},$$

et $AB \neq BA$.

Il est immédiat de vérifier que $(AB)\,C = A\,(BC)$.

Si $A$ est un opérateur linéaire, on peut interpréter la notation $Ax$ comme le produit de la matrice $A$ par la matrice colonne

$$x=\begin{Vmatrix}x_1\\ \cdots\\ x_n\end{Vmatrix}.$$

Soient donnés des opérateurs linéaires $A$ et $B$. On appelle *somme* de ces opérateurs l'opérateur $A+B$ défini par

$$(A+B)\,x = Ax + Bx, \qquad \forall x \in R_n.$$

Il est évident que la matrice de l'opérateur $A+B$ est confondue avec la matrice qui est égale à la somme des matrices des opérateurs $A$ et $B$.

Il est immédiat de vérifier que

$$A\,(B+C) = AB + AC.$$

EXEMPLE. Trouver l'inverse de la matrice

$$A=\begin{Vmatrix}1 & 2 & 0\\ 0 & 1 & 1\\ 1 & 1 & 0\end{Vmatrix}.$$

Désignons par $b_{js}$ les éléments de la matrice inverse $A^{-1}$. On a $\Delta = 1$, $b_{js} = A_{sj}$,

$$A_{11}=\begin{vmatrix}1 & 1\\ 1 & 0\end{vmatrix}=-1, \quad A_{12}=-\begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix}=1,$$

$$A_{13}=\begin{vmatrix}0 & 1\\ 1 & 1\end{vmatrix}=-1,$$

$$A_{21}=-\begin{vmatrix}2 & 0\\ 1 & 0\end{vmatrix}=0, \quad A_{22}=\begin{vmatrix}1 & 0\\ 1 & 0\end{vmatrix}=0,$$

$$A_{23}=-\begin{vmatrix}1 & 2\\ 1 & 1\end{vmatrix}=1,$$

$$A_{31} = \begin{vmatrix} 2 & 0 \\ 1 & 1 \end{vmatrix} = 2, \quad A_{32} = -\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -1,$$

$$A_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1.$$

Donc

$$\begin{aligned} &b_{11} = A_{11} = -1, \quad b_{12} = A_{21} = 0, \quad b_{13} = A_{31} = 2, \\ &b_{21} = A_{12} = 1, \quad b_{22} = A_{22} = 0, \quad b_{23} = A_{32} = -1, \\ &b_{31} = A_{13} = -1, \quad b_{32} = A_{23} = 1, \quad b_{33} = A_{33} = 1. \end{aligned}$$

Et

$$A^{-1} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

## § 16. Bases dans $R_n$

Considérons dans l'espace $R_n$ les $n$ vecteurs

$$\left\{ \begin{array}{l} i^1 = (1, 0, \ldots, 0), \\ i^2 = (0, 1, 0, \ldots, 0), \\ \cdots\cdots\cdots\cdots \\ i^n = (0, 0, \ldots, 0, 1), \end{array} \right. \tag{1}$$

dits *vecteurs unités des axes de coordonnées de* $R_n$.

On appelle *axe* $Ox_k$ de l'espace $R_n$ l'ensemble des points de la forme $(0, \ldots, 0, x_k, 0, \ldots, 0)$, où $x_k$ occupe la $k$-ième place et parcourt toutes les valeurs réelles. Le vecteur $\boldsymbol{i}^k$ est le *vecteur unité de l'axe* $Ox_k$.

Si $\boldsymbol{a} = (x_1, \ldots, x_n)$ est un vecteur arbitraire de $R_n$, on peut de toute évidence l'écrire sous forme d'une combinaison linéaire des vecteurs (1) de la manière suivante:

$$\boldsymbol{a} = x_1 \boldsymbol{i}^1 + x_2 \boldsymbol{i}^2 + \ldots + x_n \boldsymbol{i}^n. \tag{2}$$

Le système $(\boldsymbol{i}^1, \ldots, \boldsymbol{i}^n)$ est linéairement indépendant, puisque de l'égalité $\boldsymbol{a} = (x_1, \ldots, x_n) = \boldsymbol{0}$ il s'ensuit que $x_1 = \ldots = x_n = 0$.

Considérons un système de $n$ vecteurs linéairement indépendants

$$\left\{ \begin{array}{l} \boldsymbol{a}^1 = (a_{11}, \ldots, a_{1n}), \\ \cdots\cdots\cdots\cdots \\ \boldsymbol{a}^n = (a_{n1}, \ldots, a_{nn}). \end{array} \right. \tag{3}$$

On sait (voir § 14, théorème 1) que le système (3) est linéairement indépendant si le déterminant

$$\Delta = \begin{vmatrix} a_{11} \dots a_{1n} \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ a_{n1} \dots a_{nn} \end{vmatrix} \neq 0. \tag{4}$$

Si $\Delta = 0$, le système (3) est linéairement dépendant.

Le théorème 1, § 14, nous dit que $n+1$ vecteurs quelconques de $R_n$ sont linéairement dépendants, puisque le rang de la matrice formée avec leurs coordonnées est $\leqslant n$. Donc, si $\boldsymbol{a} = (x_1, \dots, x_n)$ est un vecteur arbitraire et le système de vecteurs $\boldsymbol{a}^1, \dots, \boldsymbol{a}^n$, linéairement indépendant ($\Delta \neq 0$), alors le système de vecteurs $\boldsymbol{a}^1, \dots, \boldsymbol{a}^n, \boldsymbol{a}$ est linéairement dépendant, autrement dit il existe des nombres $\lambda_1, \dots, \lambda_n, \lambda_{n+1}$ non tous nuls, tels que

$$\lambda_1 \boldsymbol{a}^1 + \dots + \lambda_n \boldsymbol{a}^n + \lambda_{n+1} \boldsymbol{a} = 0,$$

où $\lambda_{n+1} \neq 0$ (autrement le système (3) serait linéairement dépendant). D'où

$$\boldsymbol{a} = \sum_{k=1}^{n} x'_k \boldsymbol{a}^k, \tag{5}$$

où $x'_k = \frac{-\lambda_k}{\lambda_{n+1}}$ $(k = 1, \dots, n)$ sont des nombres. Exprimons la somme (5) au moyen des vecteurs unités $\boldsymbol{i}^k$ (voir (2)):

$$\boldsymbol{a} = \sum_{k=1}^{n} x'_k \left( \sum_{l=1}^{n} a_{kl} \boldsymbol{i}^l \right) = \sum_{l=1}^{n} \left( \sum_{k=1}^{n} a_{kl} x'_k \right) \boldsymbol{i}^l.$$

D'autre part, en vertu de (2) on a

$$\boldsymbol{a} = \sum_{l=1}^{n} x_l \boldsymbol{i}^l.$$

Le système $\boldsymbol{i}^1, \dots, \boldsymbol{i}^n$ étant linéairement indépendant, les coefficients en $\boldsymbol{i}^l$ sont égaux

$$x_l = \sum_{k=1}^{n} a_{kl} x'_k \quad (l = 1, \dots, n). \tag{6}$$

Donc, si les coordonnées $x_l$ du vecteur $\boldsymbol{a}$ par rapport au système $\boldsymbol{i}^1, \dots, \boldsymbol{i}^n$ sont connues, les coordonnées $x'_k$ de ce vecteur par rapport au système $\boldsymbol{a}^1, \dots, \boldsymbol{a}^n$ se déterminent à partir de (6), et ce de façon unique, puisque le déterminant du système (6) n'est pas nul.

On a prouvé que *tout vecteur $\boldsymbol{a} \in R_n$ peut être développé suivant les vecteurs de n'importe quel système linéairement indépendant $\boldsymbol{a}^1, \dots$*

. . ., $a^n$, *c'est-à-dire représenté sous la forme d'une somme* (5), *où* $x'_1$, . . ., $x'_n$ *sont des nombres définis à l'aide de* (6), *et ce de façon unique.*

Le système de vecteurs $a^1$, . . ., $a^n$ s'appelle *base* de $R_n$ en ce sens que tout vecteur $a \in R_n$ peut être représenté de façon unique par une combinaison linéaire (5) de ces vecteurs. On a démontré que *tout système linéairement indépendant de n vecteurs de $R_n$ est une base de $R_n$.*

Le système linéairement indépendant de $m$ vecteurs

$$a^1 = (a_{11}, \ldots, a_{1m}, a_{1,m+1}, \ldots, a_{1n}),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$a^m = (a_{m1}, \ldots, a_{mm}, a_{m,m+1}, \ldots, a_{mn}),$$

où $m < n$, n'est pas une base de $R_n$. En effet, le rang de la matrice des coordonnées de ces vecteurs est égal à $m$. On admettra que les $m$ premières colonnes de cette matrice forment un déterminant non nul. Elargissons cette matrice en lui ajoutant la ligne

$$a^{m+1} = (0, \ldots, 0, 1, 0, \ldots, 0),$$

où 1 se trouve au $(m + 1)$-ième rang. La matrice élargie est de rang $m + 1$ et par suite le système de vecteurs $a^1$, . . ., $a^m$, $a^{m+1}$ est linéairement indépendant. Donc, le vecteur $a^{m+1}$ ne peut être une combinaison linéaire des vecteurs du système $a^1$, . . ., $a^m$ et ce système n'est pas une base de $R_n$.

Désignons par

$$A = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix}$$

la matrice des vecteurs $(a^1, \ldots, a^n)$ [1]).

On passe de la base $(i^1, \ldots, i^n)$ à la base $(a^1, \ldots, a^n)$ à l'aide de la matrice $A$ :

$$a^k = \sum_{s=1}^{n} a_{ks} i^s \quad (k = 1, \ldots, n), \tag{7}$$

c'est-à-dire que le vecteur $a^k$ s'exprime en fonction du vecteur $i^s$ à l'aide de la $k$-ième ligne de la matrice $A$. On passe inversement de $(a^1, \ldots, a^n)$ à $(i^1, \ldots, i^n)$ à l'aide de la matrice inverse $A^{-1}$ (voir § 15, (9))

$$i^s = \sum_{l=1}^{n} b_{sl} a^l \quad (s = 1, \ldots, n), \tag{8}$$

[1]) La lecture du texte en petits caractères qui suit est à recommander, car on se refère à (9) dans le paragraphe 17.

dont les éléments sont donnés par les formules $b_{sl} = \frac{A_{ls}}{\Delta}$, où $A_{ls}$ est le cofacteur de l'élément $a_{ls}$ de $\Delta$ (signalons que l'élément $b_{sl}$ s'exprime en fonction du cofacteur $A_{ls}$ de $a_{ls}$). Notons encore que

$$a = \sum_{l=1}^{n} x'_l a^l = \sum_{s=1}^{n} x_s i^s = \sum_{s=1}^{n} x_s \sum_{l=1}^{n} b_{sl} a^l = \sum_{l=1}^{n} \Big(\sum_{s=1}^{n} b_{sl} x_s\Big) a^l,$$

d'où

$$x'_l = \sum_{s=1}^{n} b_{sl} x_s, \tag{9}$$

c'est-à-dire qu'on passe des coordonnées $(x_1, \ldots, x_n)$ aux coordonnées $(x'_1, \ldots \ldots, x'_n)$ à l'aide de la matrice

$$(A^{-1})^* = (A^*)^{-1}.$$

De (9) on voit que $x'_l$ s'exprime en fonction de $x_1, \ldots, x_n$ à l'aide de la $l$-ième colonne de la matrice $A^{-1}$ ou de la $l$-ième ligne de l'adjointe $(A^{-1})^*$ de la matrice $A^{-1}$.

Les formules (6)

$$x_s = \sum_{l=1}^{n} a_{ls} x'_l \quad (s = 1, \ldots, n)$$

nous disent qu'on passe de $(x'_1, \ldots, x'_n)$ à $(x_1, \ldots, x_n)$ à l'aide de l'adjointe $A^*$ de la matrice $A$, c'est-à-dire que $x_s$ s'exprime en fonction de $x'_1, \ldots, x'_n$ à l'aide de la $s$-ième ligne de la matrice $A^*$ ou de la $s$-ième colonne de la matrice $A$.

REMARQUE. Au § 14 on a vu qu'une matrice carrée arbitraire

$$A = \begin{Vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{Vmatrix} \tag{10}$$

définissait un opérateur linéaire $y = Ax$ $(x \in R_n,\ y \in R_n)$ donné par

$$y_k = \sum_{j=1}^{n} a_{kj} x_j \quad (k = 1, \ldots, n). \tag{11}$$

La réciproque est aussi vraie: tout opérateur linéaire $y = Ax$ $(x \in R_n,\ y \in R_n)$ est défini par une matrice (10) telle que le vecteur $y = Ax$ s'exprime en fonction du vecteur $x$ par les formules (11).

En effet, soit donné un opérateur linéaire quelconque $y = Ax$ $(x \in R_n,\ y \in R_n)$. Représentons les images des vecteurs unités $i^s$ par $A$ de la manière suivante:

$$A(i^s) = (a_{1s}, a_{2s}, \ldots, a_{ns}) = \sum_{k=1}^{n} a_{ks} i^k \quad (s = 1, \ldots, n).$$

En vertu de la linéarité de $A$, l'image de tout vecteur

$$x = x_1 i^1 + \ldots + x_n i^n = \sum_{s=1}^{n} x_s i^s$$

par $A$ est un vecteur $y$ défini par

$$y = Ax = A\Big(\sum_{s=1}^{n} x_s i^s\Big) = \sum_{s=1}^{n} x_s A(i^s) = \sum_{s=1}^{n} x_s \sum_{k=1}^{n} a_{ks} i^k = \sum_{k=1}^{n} \Big(\sum_{s=1}^{n} a_{ks} x_s\Big) i^k,$$

d'où il résulte que la $k$-ième composante de $y$ est déterminée par la formule (11). Donc, l'opérateur $A$ engendre une matrice (10) dont les éléments des colonnes sont les coordonnées des images des vecteurs unités par $A$.

EXEMPLE. Trouver la matrice de l'opérateur linéaire (transformation) $A$ qui consiste en une rotation des vecteurs de $R_2$, issus de l'origine, d'un angle $\alpha$ $(0 < \alpha < \pi/2)$ dans le sens rétrograde.

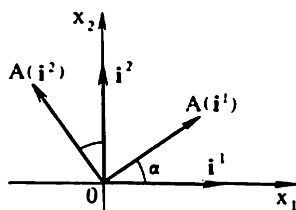



Fig. 142

Prenons les vecteurs $i^1 = (1, 0)$ et $i^2 = (0, 1)$ pour base. On a alors (fig. 142)

$$A\ (i^1) = (\cos\alpha,\ \sin\alpha),$$
$$A\ (i^2) = (-\sin\alpha,\ \cos\alpha).$$

Donc, la matrice cherchée s'écrit

$$A = \begin{Vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{Vmatrix}.$$

## § 17. Bases orthogonales de $R_n$

On dit que deux vecteurs non nuls $x$ et $y$ de $R_n$ sont de *même sens* s'il existe un nombre $\lambda > 0$ tel que $x = \lambda y$.

Tout vecteur $x \in R_n$ non nul peut être *normé*, c'est-à-dire remplacé par un vecteur unité

$$y = \frac{1}{|x|} x \quad (|y| = 1)$$

de même sens.

On dit que deux vecteurs $x$ et $y$ de $R_n$ sont *orthogonaux* si leur produit scalaire est nul: $(x, y) = 0$.

On dit qu'un système de vecteurs

$$x^1, \ldots, x^\nu \in R_n \tag{1}$$

est *orthogonal* si ses vecteurs sont deux à deux orthogonaux. Le système (1) est *orthonormal* (ou *orthonormé*) si

$$(x^k, x^l) = \delta_{kl} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases} \quad (k, l = 1, \ldots, \nu),$$

c'est-à-dire que tous les vecteurs de ce système sont des vecteurs unités deux à deux orthogonaux. Si le système de vecteurs (1) est orthogonal et si aucun d'eux n'est nul, en les normant on obtient manifestement un système orthonormal. *Le système orthonormal* (1) *est linéairement indépendant*. En effet, soit

$$\lambda_1 x^1 + \lambda_2 x^2 + \ldots + \lambda_\nu x^\nu = 0,$$

où $\lambda_1, \ldots, \lambda_\nu$ sont des nombres. En multipliant cette égalité scalairement par $x^s$ on obtient visiblement

$$\lambda_s (x^s, x^s) = \lambda_s = 0 \quad (s = 1, \ldots, \nu).$$

Donc, un système orthonormal de $n$ vecteurs est une base de $R_n$ et par suite tout vecteur $a \in R_n$ peut être représenté par une combinaison linéaire de la forme

$$a = \sum_{k=1}^{n} \lambda_k x^k. \tag{2}$$

En multipliant cette égalité scalairement par $x^s$, on obtient

$$(a, x^s) = \lambda_s \quad (s = 1, \ldots, n)$$

et par suite

$$a = \sum_{k=1}^{n} (a, x^k) x^k, \quad \forall a \in R_n.$$

Le nombre $(a, x^k)$ $(|x^k| = 1!)$ s'appelle *projection du vecteur* $a$ sur la direction du vecteur $x^k$.

THÉORÈME. *Un système orthonormal de vecteurs*

$$x^1, x^2, \ldots, x^\nu \quad (\nu < n)$$

*peut être complété à une base orthonormale de* $R_n$. *Autrement dit, on peut exhiber des vecteurs* $x^{\nu+1}, \ldots, x^n$, *tels que le système*

$$x^1, \ldots, x^\nu, x^{\nu+1}, \ldots, x^n \tag{3}$$

*soit orthonormal et par suite soit une base de* $R_n$.

DÉMONSTRATION. Comme $\nu < n$, il existe dans $R_n$ un vecteur $a$ linéairement indépendant de $x^1, \ldots, x^\nu$. Alors

$$a = \sum_{k=1}^{\nu} (a, x^k) x^k + y,$$

où $y \neq 0$. Le vecteur $y$ est orthogonal à tous les vecteurs $x^1, \ldots, x^\nu$. En effet

$$(y, x^s) = \left(a - \sum_{k=1}^{\nu} (a, x^k) x^k, x^s\right) = (a, x^s) - (a, x^s) = 0 \tag{4}$$

$$(s = 1, \ldots, \nu).$$

En normant $y$ on obtient le vecteur

$$x^{\nu+1}=\frac{1}{|y|}y \quad (|x^{\nu+1}|=1),$$

et le système

$$x^1, \ldots, x^\nu, x^{\nu+1}$$

sera orthonormal. Si $\nu+1=n$, on obtient une base de $R_n$, sinon on poursuit cette procédure. Au $(n-\nu)$-ième pas on obtient la base (3) de $R_n$.

Le système des vecteurs unités des axes de $R_n$:

$$\begin{aligned} i^1 &= (1,\ 0,\ 0,\ \ldots,\ 0,\ 0),\\ i^2 &= (0,\ 1,\ 0,\ \ldots,\ 0,\ 0),\\ &\ldots\ldots\ldots\ldots\ldots\\ i^n &= (0,\ 0,\ 0,\ \ldots,\ 0,\ 1) \end{aligned}$$

nous fournit un exemple de base orthonormale de $R_n$.

Tout vecteur $a=(x_1, \ldots, x_n)\in R_n$ se décompose suivant les vecteurs unités de la manière suivante:

$$a=x_1i^1+\ldots+x_ni^n=\sum_{k=1}^{n} x_ki^k, \tag{5}$$

où $x_k=(a, i^k)$ $(k=1, \ldots, n)$ est la projection du vecteur $a$ sur la direction du vecteur unité $i^k$.

Soit donné un système orthonormal de $n$ vecteurs:

$$\left\{\begin{array}{l} a^1=(a_{11}, \ldots, a_{1n}),\\ \ldots\ldots\ldots\ldots\\ a^n=(a_{n1}, \ldots, a_{nn}), \end{array}\right. \tag{6}$$

i.e.

$$a^k=\sum_{s=1}^{n} a_{ks}i^s \quad (k=1, \ldots, n). \tag{7}$$

On passe des vecteurs $(i^1, \ldots, i^n)$ aux vecteurs $(a^1, \ldots, a^n)$ à l'aide de la matrice

$$\Lambda=\left\|\begin{array}{ccc} a_{11} & \ldots & a_{1n}\\ \cdot & \cdots & \cdot\\ a_{n1} & \ldots & a_{nn} \end{array}\right\|, \tag{8}$$

autrement dit, le vecteur $a^k$ s'exprime en fonction de $i^1, \ldots, i^n$ à l'aide des éléments de la $k$-ième ligne de la matrice $\Lambda$.

*La matrice* $\Lambda$ *est orthogonale*, c'est-à-dire possède la propriété suivante:

$$\sum_{s=1}^{n} a_{ks}a_{ls}=\delta_{kl} \quad (k,\ l=1, \ldots, n). \tag{9}$$

En effet, le système $a^1, \ldots, a^n$ étant orthonormal, on a

$$\delta_{kl} = (a^k, a^l) = \left(\sum_{s=1}^{n} a_{ks} i^s, \sum_{r=1}^{n} a_{lr} i^r\right) =$$
$$= \sum_{s=1}^{n} \sum_{r=1}^{n} a_{ks} a_{lr} (i^s, i^r) = \sum_{s=1}^{n} \sum_{r=1}^{n} a_{ks} a_{lr} \delta_{sr} = \sum_{s=1}^{n} a_{ks} a_{ls}. \quad (10)$$

On voit que réciproquement l'orthogonalité de la matrice (8) entraîne l'orthonormalité du système de vecteurs $a^1, \ldots, a^n$, définis par (7).

Ceci montre que *les formules* (7), *où* $\| a_{ks} \|$ *sont des matrices orthogonales quelconques, définissent toutes les bases orthonormales de* $R_n$.

Multiplions l'équation (7) scalairement par $i^l$:

$$(a^k, i^l) = a_{kl}.$$

D'où

$$i^l = \sum_{k=1}^{n} (a^k, i^l) a^k = \sum_{k=1}^{n} a_{kl} a^k. \quad (11)$$

Donc, on passe de la base $(a^1, \ldots, a^n)$ à la base $(i^1, \ldots, i^n)$ à l'aide de la matrice $\Lambda^*$ adjointe de $\Lambda$. Comme les transformations (11) sont inverses des transformations (7) (voir § 15), on a démontré incidemment que la matrice orthogonale $\Lambda$ est douée de la remarquable propriété suivante:

$$\Lambda^{-1} = \Lambda^*.$$

De (11) il s'ensuit que

$$\delta_{kl} = (i^k, i^l) = \left(\sum_{s=1}^{n} a_{sk} a^s, \sum_{r=1}^{n} a_{rl} a^r\right) =$$
$$= \sum_{s=1}^{n} \sum_{r=1}^{n} a_{sk} a_{rl} (a^s, a^r) = \sum_{s=1}^{n} a_{sk} a_{sl}. \quad (12)$$

Nous avons défini une matrice orthogonale comme une matrice dont les vecteurs lignes sont des vecteurs unités deux à deux orthogonaux. De cette définition il s'ensuit automatiquement que les vecteurs colonnes d'une matrice orthogonale sont aussi des vecteurs unités deux à deux orthogonaux (voir (12)).

On passe de $(x_1, \ldots, x_n)$ à $(x'_1, \ldots, x'_n)$ à l'aide de la matrice (voir (9), § 16, $x' = \Lambda x$)

$$(\Lambda^*)^{-1} = \Lambda^{**} = \Lambda,$$

c'est-à-dire que (en admettant que $a = \sum_{l=1}^{n} x'_l a^l$)

$$x'_l = \sum_{s=1}^{n} a_{ls} x_s. \quad (13)$$

On passe inversement de $(x'_1, \ldots, x'_n)$ à $(x_1, \ldots, x_n)$ à l'aide de la matrice $\Lambda^*$ adjointe de $\Lambda$ (voir (6), § 16), c'est-à-dire que

$$x_s = \sum_{l=1}^{n} a_{ls} x'_l.$$

Signalons que le déterminant d'une matrice orthogonale (voir (6)) est égal à 1 en module: $\| \Lambda \| = \| a_{kl} \| = 1$.

Ceci résulte de ce que

$$|\Lambda|^2 = |a_{kl}| \cdot |a_{ij}| = \left| \sum_{s=1}^{n} a_{ks} a_{ls} \right| = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} = 1.$$

On admet ici que l'élément $\gamma_{kl}$ du produit des déterminants est égal à la somme des produits des éléments de la $k$-ième ligne par les éléments correspondants de la $i$-ième ligne (voir § 2, propriété $j$)).

Signalons encore que le déterminant formé avec les coordonnées des vecteurs de la base $\boldsymbol{i}^1, \ldots, \boldsymbol{i}^n$ est égal à 1 :

$$\begin{vmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & 1 \end{vmatrix} = 1.$$

Si une base orthogonale $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^n$ possède un déterminant $|\Lambda| = 1$ (voir (6)), on dit qu'*elle est orientée comme la base* $\boldsymbol{i}^1, \ldots, \boldsymbol{i}^n$. Si $|\Lambda| = -1$, elle est orientée *dans le sens inverse*. Ces définitions sont compatibles avec celles énoncées au § 11 pour les dimensions deux et trois.

REMARQUE. Si l'on avait porté les coordonnées du vecteur $\boldsymbol{a}^k$ relatives à la base $(\boldsymbol{i}^1, \ldots, \boldsymbol{i}^n)$ dans la $k$-ième colonne de la matrice $\Lambda$ (voir (8)), on serait passé des coordonnées du vecteur $\boldsymbol{x}'$ à celles du vecteur $\boldsymbol{x}$ à l'aide des lignes de la matrice $\Lambda$. Dans la formule (13), on serait passé de $\boldsymbol{x}$ à $\boldsymbol{x}'$ à l'aide de la matrice $\Lambda^*$.

## § 18. Propriétés invariantes des produits scalaire et vectoriel

REMARQUE. Soient donnés dans $R_3$ deux systèmes de coordonnées rectangulaires $(x_1, x_2, x_3)$ et $(y_1, y_2, y_3)$ de vecteurs unités respectifs $(\boldsymbol{i}^1, \boldsymbol{i}^2, \boldsymbol{i}^3)$ et $(\boldsymbol{j}^1, \boldsymbol{j}^2, \boldsymbol{j}^3)$. Soit

$$\boldsymbol{j}^k = \sum_{s=1}^{3} \alpha_{ks} \boldsymbol{i}^s \quad (k = 1, 2, 3) \tag{1}$$

(cf. (7) et (13), § 17). Alors la matrice

$$A = \| \alpha_{ks} \|$$

est orthogonale et (cf. (13), § 17)

$$y_l = \sum_{s=1}^{3} \alpha_{ls} x_s \quad (l = 1, 2, 3). \tag{2}$$

Un même point de l'espace possède les coordonnées $(x_1, x_2, x_3)$ dans le premier système et $(y_1, y_2, y_3)$ dans le second. Les formules de changement des coordonnées (c'est-à-dire le passage de $(x_1, x_2, x_3)$ à $(y_1, y_2, y_3)$) sont exactement les mêmes que les formules de passage du système de vecteurs unités $(i^1, i^2, i^3)$ au système $(j^1, j^2, j^3)$. Dans les deux cas on se sert de la même matrice orthogonale (voir (13) et (7), § 17)

$$A = \| \alpha_{kl} \|,$$

dans le premier, pour passer du système de nombres $(x_1, x_2, x_3)$ au système $(y_1, y_2, y_3)$, dans le second, pour passer du système $(i^1, i^2, i^3)$ au système $(j^1, j^2, j^3)$.

Voici la définition générale d'un vecteur en calcul tensoriel.

*On appelle vecteur de $R_3$ un être exprimé dans chaque système de coordonnées rectangulaires par un triplet de nombres qui se transforme de la même manière (et à l'aide de la même matrice) que les triplets de vecteurs unités des systèmes de coordonnées respectifs.*

Cette terminologie est également en usage dans $R_n$.

Soient $\boldsymbol{a}$ et $\boldsymbol{b}$ des vecteurs de $R_3$ définis dans des systèmes de coordonnées de vecteurs unités $(i^1, i^2, i^3)$ et $(j^1, j^2, j^3)$ de la manière suivante:

$$\boldsymbol{a} = a_{x_1} i^1 + a_{x_2} i^2 + a_{x_3} i^3 = a_{y_1} j^1 + a_{y_2} j^2 + a_{y_3} j^3,$$
$$\boldsymbol{b} = b_{x_1} i^1 + b_{x_2} i^2 + b_{x_3} i^3 = b_{y_1} j^1 + b_{y_2} j^2 + b_{y_3} j^3.$$

On a l'égalité

$$(\boldsymbol{a}, \boldsymbol{b}) = \sum_{s=1}^{3} a_{x_s} b_{x_s} = \sum_{r=1}^{3} a_{y_r} b_{y_r},$$

qui dit que *le produit scalaire est invariant par les changements des systèmes de coordonnées rectangulaires.*

En effet, les systèmes $(i^1, i^2, i^3)$ et $(j^1, j^2, j^3)$ étant orthonormaux, l'on passe de l'un à l'autre à l'aide des formules (1). où $\| \alpha_{kl} \|$ est une matrice orthogonale. Les coordonnées $(a_{x_1}, a_{x_2}, a_{x_3})$ du vecteur $\boldsymbol{a}$ se transforment en les coordonnées $(a_{y_1}, a_{y_2}, a_{y_3})$ à l'aide de la même matrice (voir (2)). Donc

$$\sum_{r=1}^{3} a_{y_r} a_{y_r} = \sum_{r=1}^{3} \Big( \sum_{s=1}^{3} \alpha_{rs} a_{x_s} \cdot \sum_{\nu=1}^{3} \alpha_{r\nu} b_{x_\nu} \Big) =$$
$$= \sum_{s=1}^{3} \sum_{\nu=1}^{3} a_{x_s} b_{x_\nu} \Big( \sum_{r=1}^{3} \alpha_{rs} \alpha_{r\nu} \Big) = \sum_{s=1}^{3} a_{x_s} b_{x_s} = (\boldsymbol{a}, \boldsymbol{b}). \tag{3}$$

Nous avons montré l'invariance du produit scalaire $(\boldsymbol{a}, \boldsymbol{b})$ par voie de calcul. Du reste, de la définition géométrique du produit scalaire des vecteurs qui dit que $(\boldsymbol{a}, \boldsymbol{b}) = |\boldsymbol{b}| \cdot \mathrm{pr}_{\boldsymbol{b}}\, \boldsymbol{a}$ il s'ensuit immédiatement que ce nombre est un invariant, puisque cette définition est intrinsèque, c'est-à-dire n'est liée à aucun système de coordonnées.

S'agissant du produit vectoriel $\boldsymbol{a} \times \boldsymbol{b}$, il est invariant par tout changement des systèmes de coordonnées rectangulaires conservant l'orientation. Dans les systèmes de vecteurs unités $(\boldsymbol{i}^1, \boldsymbol{i}^2, \boldsymbol{i}^3)$ et $(\boldsymbol{j}^1, \boldsymbol{j}^2, \boldsymbol{j}^3)$, les produits vectoriels s'expriment comme suit:

$$[\boldsymbol{a} \times \boldsymbol{b}]_{\boldsymbol{i}^1, \boldsymbol{i}^2, \boldsymbol{i}^3} = \begin{vmatrix} \boldsymbol{i}^1 & \boldsymbol{i}^2 & \boldsymbol{i}^3 \\ a_{x_1} & a_{x_2} & a_{x_3} \\ b_{x_1} & b_{x_2} & b_{x_3} \end{vmatrix},$$

$$[\boldsymbol{a} \times \boldsymbol{b}]_{\boldsymbol{j}^1, \boldsymbol{j}^2, \boldsymbol{j}^3} = \begin{vmatrix} \boldsymbol{j}^1 & \boldsymbol{j}^2 & \boldsymbol{j}^3 \\ a_{y_1} & a_{y_2} & a_{y_3} \\ b_{y_1} & b_{y_2} & b_{y_3} \end{vmatrix}.$$

En vertu des formules (1) et (2)

$$[\boldsymbol{a} \times \boldsymbol{b}]_{\boldsymbol{j}^1, \boldsymbol{j}^2, \boldsymbol{j}^3} = \begin{vmatrix} \sum\limits_{s=1}^{3} \alpha_{1s}\boldsymbol{i}^s & \sum\limits_{s=1}^{3} \alpha_{2s}\boldsymbol{i}^s & \sum\limits_{s=1}^{3} \alpha_{3s}\boldsymbol{i}^s \\ \sum\limits_{s=1}^{3} \alpha_{1s}a_{x_s} & \sum\limits_{s=1}^{3} \alpha_{2s}a_{x_s} & \sum\limits_{s=1}^{3} \alpha_{3s}a_{x_s} \\ \sum\limits_{s=1}^{3} \alpha_{1s}b_{x_s} & \sum\limits_{s=1}^{3} \alpha_{2s}b_{x_s} & \sum\limits_{s=1}^{3} \alpha_{3s}b_{x_s} \end{vmatrix} =$$

$$= \begin{vmatrix} \boldsymbol{i}^1 & \boldsymbol{i}^2 & \boldsymbol{i}^3 \\ a_{x_1} & a_{x_2} & a_{x_3} \\ b_{x_1} & b_{x_2} & b_{x_3} \end{vmatrix} \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = \pm \begin{vmatrix} \boldsymbol{i}^1 & \boldsymbol{i}^2 & \boldsymbol{i}^3 \\ a_{x_1} & a_{x_2} & a_{x_3} \\ b_{x_1} & b_{x_2} & b_{x_3} \end{vmatrix} = \pm [\boldsymbol{a} \times \boldsymbol{b}]_{\boldsymbol{i}^1, \boldsymbol{i}^2, \boldsymbol{i}^3}, \quad (4)$$

où l'on prend le signe + ou — selon que le déterminant $|\alpha_{ik}|$ est égal à +1 ou —1, ou ce qui revient au même selon que le changement de coordonnées modifie ou non l'orientation.

Lors du produit des déterminants on s'est servi de la règle suivante: l'élément $\gamma_{ik}$ de la matrice du produit $\|\gamma_{ik}\|$ est le produit de la $i$-ième ligne du premier déterminant par la $k$-ième colonne du second (voir § 2, propriété $j$)).

Ainsi, nous avons montré par des calculs que *le produit vectoriel de deux vecteurs est invariant par les changements des systèmes de coordonnées rectangulaires conservant l'orientation.*

Les formules (3) et (4) sont intéressantes par ce qu'elles se généralisent au cas où $\boldsymbol{a}$ est le vecteur symbolique $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)$ qui joue un rôle fondamental en analyse.

29—0622

## § 19. Changement des systèmes de coordonnées rectangulaires dans le plan

Considérons un plan $R_2$ rapporté à un système de coordonnées rectangulaires $(x_1, x_2)$. Soient

$$i^1 = (1, 0), \quad i^2 = (0, 1)$$

fes vecteurs unités des axes $Ox_1$ et $Ox_2$. Les vecteurs unités $i^1$ et $i^2$ lorment une base orthonormale de $R_2$.

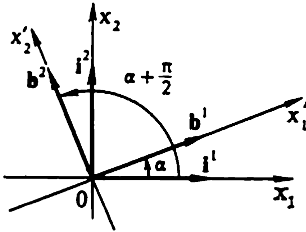



Fig. 143

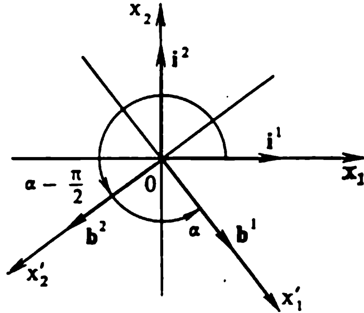



Fig. 144

Tout vecteur unité $b^1$ peut être mis sous la forme:

$$b = (\cos \alpha, \sin \alpha) \quad (0 \leqslant \alpha < 2\pi).$$

Le vecteur unité orthogonal à $b^1$, vecteur que nous désignerons par $b^2$, fait avec l'axe $Ox_1$ un angle de $\alpha + \frac{\pi}{2}$ ou de $\alpha - \frac{\pi}{2}$. Comme

$$\cos\left(\alpha + \frac{\pi}{2}\right) = -\sin \alpha, \quad \sin\left(\alpha + \frac{\pi}{2}\right) = \cos \alpha,$$

$$\cos\left(\alpha - \frac{\pi}{2}\right) = \sin \alpha, \quad \sin\left(\alpha - \frac{\pi}{2}\right) = -\cos \alpha,$$

les systèmes orthonormaux $(b^1, b^2)$ de $R_2$ sont tous définis soit par les égalités (fig. 143)

$$\left\{\begin{array}{l} b^1 = i^1 \cos \alpha + i^2 \sin \alpha, \\ b^2 = -i^1 \sin \alpha + i^2 \cos \alpha \end{array}\right. \quad (0 \leqslant \alpha < 2\pi) \qquad (1')$$

qui correspondent à une rotation d'un angle $\alpha$ autour de l'origine des coordonnées sans changement de l'orientation, soit par les égalités (fig. 144)

$$\left\{\begin{array}{l} b^1 = i^1 \cos \alpha + i^2 \sin \alpha, \\ b^2 = i^1 \sin \alpha - i^2 \cos \alpha \end{array}\right. \qquad (1'')$$

qui correspondent à une rotation d'un angle $\alpha$ autour de l'origine des coordonnées avec changement de l'orientation.

Ces deux formules de changement se combinent en la formule suivante :

$$\begin{aligned} \boldsymbol{b}^1 &= \alpha_{11}\boldsymbol{i}^1 + \alpha_{12}\boldsymbol{i}^2, \\ \boldsymbol{b}^2 &= \alpha_{21}\boldsymbol{i}^1 + \alpha_{22}\boldsymbol{i}^2, \end{aligned} \tag{1}$$

où la matrice

$$\Lambda = \begin{Vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{Vmatrix} \tag{2}$$

est orthogonale (la somme des carrés des éléments de toute ligne ou de toute colonne est égale à 1, quant au produit scalaire de deux lignes ou de deux colonnes différentes il est nul).

Toute transformation orthogonale (1) est en fait une transformation (1′) ou (1″) pour un certain $\alpha$.

De (1) il s'ensuit, compte tenu de l'orthogonalité de la matrice (2), que

$$\left\{ \begin{aligned} \boldsymbol{i}^1 &= \alpha_{11}\boldsymbol{b}^1 + \alpha_{21}\boldsymbol{b}^2, \\ \boldsymbol{i}^2 &= \alpha_{12}\boldsymbol{b}^1 + \alpha_{22}\boldsymbol{b}^2, \end{aligned} \right. \tag{3}$$

et l'on obtient la transformation inverse de (1), dont la matrice

$$\begin{Vmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{Vmatrix} = \Lambda^*$$

est adjointe de $\Lambda$.

Soit un vecteur arbitraire $\boldsymbol{a} \in R_2$. Supposons que ses coordonnées respectives dans l'ancien et le nouveau système de coordonnées sont $(x_1, x_2)$ et $(x_1', x_2')$. Alors

$$\boldsymbol{a} = x_1\boldsymbol{i}^1 + x_2\boldsymbol{i}^2 = x_1'\boldsymbol{b}^1 + x_2'\boldsymbol{b}^2. \tag{4}$$

Les formules (3) et (4) nous donnent

$$\begin{aligned} x_1'\boldsymbol{b}^1 + x_2'\boldsymbol{b}^2 &= x_1(\alpha_{11}\boldsymbol{b}^1 + \alpha_{21}\boldsymbol{b}^2) + x_2(\alpha_{12}\boldsymbol{b}^1 + \alpha_{22}\boldsymbol{b}^2) = \\ &= (\alpha_{11}x_1 + \alpha_{12}x_2)\,\boldsymbol{b}^1 + (\alpha_{21}x_1 + \alpha_{22}x_2)\,\boldsymbol{b}^2. \end{aligned}$$

En identifiant les termes des deux membres on obtient

$$\left\{ \begin{aligned} x_1' &= \alpha_{11}x_1 + \alpha_{12}x_2, \\ x_2' &= \alpha_{21}x_1 + \alpha_{22}x_2. \end{aligned} \right. \tag{5}$$

Les formules (1) et (4) nous donnent

$$\begin{aligned} x_1\boldsymbol{i}^1 + x_2\boldsymbol{i}^2 &= x_1'(\alpha_{11}\boldsymbol{i}^1 + \alpha_{12}\boldsymbol{i}^2) + x_2'(\alpha_{21}\boldsymbol{i}^1 + \alpha_{22}\boldsymbol{i}^2) = \\ &= (\alpha_{11}x_1' + \alpha_{21}x_2')\,\boldsymbol{i}^1 + (\alpha_{12}x_1' + \alpha_{22}x_2')\,\boldsymbol{i}^2, \end{aligned}$$

d'où par identification des coefficients en $i^1$ et $i^2$, l'on obtient les formules inverses de (5):

$$\begin{cases} x_1 = \alpha_{11}x_1' + \alpha_{21}x_2', \\ x_2 = \alpha_{12}x_1' + \alpha_{22}x_2'. \end{cases} \tag{6}$$

Si en plus du changement (6) on transporte l'origine des coordonnées en un point $O' = (x_1^0, x_2^0)$, alors les formules (6) se compliquent et deviennent

$$\begin{cases} x_1 = x_1^0 + \alpha_{11}x_1^1 + \alpha_{21}x_2^1, \\ x_2 = x_2^0 + \alpha_{12}x_1' + \alpha_{22}x_2'. \end{cases} \tag{7}$$

Ainsi, tout passage des coordonnées $(x_1, x_2)$ aux coordonnées $(x_1', x_2')$ avec translation de l'origine du système $(x_1, x_2)$ en un point $O' = (x_1^0, x_2^0)$ s'exprime par les formules (7), où la matrice

$$\begin{Vmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{Vmatrix}$$

est orthogonale.

Les formules de passage des anciennes coordonnées aux nouvelles s'écrivent

$$\begin{cases} x_1 = x_1^0 + x_1' \cos\alpha - x_2' \sin\alpha, \\ x_2 = x_2^0 + x_1' \sin\alpha + x_2' \cos\alpha \end{cases} \tag{7'}$$

si l'orientation est conservée, et

$$\begin{cases} x_1 = x_1^0 + x_1' \cos\alpha + x_2' \sin\alpha, \\ x_2 = x_2^0 + x_1' \sin\alpha - x_2' \cos\alpha \end{cases} \tag{7''}$$

si l'orientation est modifiée (les matrices des coefficients en $x_1'$ et $x_2'$ des formules (7′) et (7″) sont les transposées respectives des matrices de (1′) et (1″)).

## § 20. Sous-espaces vectoriels de $R_n$

Un ensemble $L$ de $R_n$ ($L \subset R_n$) s'appelle *sous-espace vectoriel de l'espace* $R_n$ ou simplement sous-espace de $R_n$ si l'appartenance de deux vecteurs quelconques $\boldsymbol{x}$ et $\boldsymbol{y}$ à $L$ ($\boldsymbol{x}, \boldsymbol{y} \in L$) entraîne automatiquement l'appartenance du vecteur $\alpha\boldsymbol{x} + \beta\boldsymbol{y}$ à $L$ ($\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in L$), où $\alpha$ et $\beta$ sont des nombres. Un *sous-espace* $L$ est *à* $m$ *dimensions* s'il contient un système de $m$ vecteurs $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m$ linéairement indépendants et aucun système de $m+1$ vecteurs linéairement indépendants.

Donc, si $\boldsymbol{a}$ est un vecteur de $L$, alors le système $\boldsymbol{a}^1, \ldots, \boldsymbol{a}^m, \boldsymbol{a}$ est linéairement dépendant, autrement dit il existe un système de nombres non tous nuls $\lambda_1, \ldots, \lambda_m, \lambda_{m+1}$, tel que

$$\lambda_1\boldsymbol{a}^1 + \ldots + \lambda_m\boldsymbol{a}^m + \lambda_{m+1}\boldsymbol{a} = 0. \tag{1}$$

Le nombre $\lambda_{m+1} \neq 0$, sinon on aurait

$$\lambda_1 a^1 + \ldots + \lambda_m a^m = 0,$$

et $\lambda_1 = \ldots = \lambda_m = 0$ par suite de l'indépendance linéaire du système $a^1, \ldots, a^m$. Donc, l'équation (1) est résoluble en $a$:

$$a = \mu_1 a^1 + \ldots + \mu_m a^m \quad (\mu_s = -\lambda_s/\lambda_{m+1}), \tag{2}$$

c'est-à-dire représentable par une combinaison linéaire des vecteurs $a^1, \ldots, a^m$. Par ailleurs, toute combinaison linéaire de la forme (2) appartient à $L$, puisque $L$ est un sous-espace. En ce sens, on dit que le système $a^1, \ldots, a^m$ est une *base* de $L$. Il est évident que tout autre système de vecteurs linéairement indépendants $b^1, \ldots, b^m$ de $L$ est une base de $L$.

Si l'on développe les vecteurs $b^i$ suivant les vecteurs $a^1, \ldots$ $\ldots, a^m$, on obtient

$$b^k = \sum_{s=1}^{m} b_{ks} a^s \quad (k = 1, \ldots, m).$$

Par les mêmes raisonnements qu'au § 16 pour $R_n$ (mais en changeant $i^s$ et $a^k$ respectivement par $a^s$ et $b^k$) on démontre que le système $b^1, \ldots, b^m$ est linéairement indépendant si et seulement si le déterminant $| b_{ks} | \neq 0$ et que tout système indépendant composé de $l$ ($l < m$) vecteurs ne peut être une base de $L$.

L'espace $R_n$ peut être considéré comme un sous-espace de $R_n$ à $n$ dimensions.

L'ensemble constitué du seul vecteur $\mathbf{0}$ est un sous-espace vectoriel ($\alpha\mathbf{0} + \beta\mathbf{0} = \mathbf{0}$). On dit qu'il est de dimension zéro. Le vecteur $\mathbf{0}$ ne forme pas un système linéairement indépendant: de l'égalité $\lambda\mathbf{0} = \mathbf{0}$, où $\lambda$ est un nombre, il ne s'ensuit pas forcément que $\lambda$ est nul.

Si un vecteur $x^0 \neq \mathbf{0}$, l'ensemble des vecteurs de la forme $\lambda x^0$, où $\lambda$ est un nombre arbitraire, est un *sous-espace à une dimension*. Pour base de cet espace, on peut prendre le vecteur $x^0$.

Soit $L$ un sous-espace de $R_n$. On dira qu'un vecteur $v \in R_n$ *est orthogonal à* $L$ s'il l'est à tout vecteur $u \in L$. Soit $L'$ l'ensemble de tous les vecteurs orthogonaux à $L$. L'ensemble $L'$ est un sous-espace de $R_n$. En effet, soient $v, v' \in L'$. On a

$$(v, u) = 0, \quad \forall u \in L;$$

$$(v', u) = 0, \quad \forall u \in L.$$

Alors, pour tous nombres $\alpha, \beta$

$$(\alpha v + \beta v', u) = \alpha (v, u) + \beta (v', u) = 0, \quad \forall u \in L,$$

i.e. $\alpha v + \beta v' \in L'$.

Par définition, *un sous-espace* $L' \subset R_n$ *est orthogonal à un sous-espace donné* $L \subset R_n$ si $L'$ est l'ensemble de *tous* les vecteurs orthogonaux à $L$.

On démontre plus bas un théorème qui met en lumière la structure de tout sous-espace $L \subset R_n$ et du sous-espace $L' \subset R_n$ qui lui est orthogonal. De ce théorème il s'ensuit en particulier que *si $L'$ est orthogonal à $L$, alors inversement $L$ est orthogonal à $L'$.*

THÉORÈME 1. *Soit $L$ un sous-espace différent de $R_n$ et du sous-espace nul. Alors:*

a) *il existe un entier $m$ vérifiant la relation*

$$1 \leqslant m < n, \tag{3}$$

*et une base orthonormale*

$$a^1, \ldots, a^m \tag{4}$$

*dans $L$; si l'on prolonge cette base par un procédé quelconque à une base orthonormale de $R_n$*

$$a^1, \ldots, a^m, a^{m+1}, \ldots, a^n, \tag{5}$$

*alors le sous-espace $L'$ muni de la base*

$$a^{m+1}, \ldots, a^n \tag{6}$$

*possède les propriétés suivantes:*

b) *$L'$ est orthogonal à $L$;*

c) *$L$ est orthogonal à $L'$;*

d) *tout vecteur $a \in R_n$ se représente par la somme*

$$a = u + v,$$

*où $u \in L$, $v \in L'$, et ce de façon unique.*

DÉMONSTRATION. Par hypothèse, $L$ n'est pas un sous-espace nul, donc il existe dans $L$ un vecteur $x$ non nul. En normant ce vecteur, on obtient le vecteur unité

$$a^1 = \frac{1}{|x|} x.$$

Désignons par $a^2$ un vecteur unité (s'il existe) de $L$ orthogonal à $a^1$ ($|a^2| = 1$, $(a^2, a^1) = 0$). Désignons maintenant par $a^3$ un vecteur unité (s'il existe) de $L$ orthogonal à $a^1$ et à $a^2$ ($|a^3| = 1$, $(a^3, a^1) = (a^3, a^2) = 0$). Ce processus prend fin en $m$ pas, où $m$ vérifie la relation (3), autrement dit il existe un système orthonormal de vecteurs (4) appartenant à $L$, et $L$ ne contient aucun vecteur unité orthogonal à $a^1, \ldots, a^m$. En effet, $m \geqslant 1$ car à fortiori $a^1 \in L$. D'autre part, $m \neq n$. Si $m = n$, les vecteurs $a^1, \ldots, a^n$ appartiendraient à $L$ avec toutes les combinaisons linéaires $\sum_{k=1}^{n} \lambda_k a^k$, et $L$ serait confondu avec $R_n$, ce qui est contraire à l'hypothèse. Le système orthonormal $a^1, \ldots, a^m$ obtenu est une base de $L$. En effet, les vecteurs $a^1, \ldots, a^m$ appartiennent à $L$ avec toutes leurs combi-

naisons linéaires $\sum_{k=1}^{m} \lambda_k a^k$. Mais $L$ ne contient pas d'autres vecteurs, car si l'on admet qu'il existe un vecteur $a \in L$ qui ne soit pas une telle combinaison linéaire, alors $a$ pourrait être représenté par une somme

$$a = \sum_{k=1}^{m} (a, a^k) a^k + y, \tag{7}$$

où $y \neq 0$. Comme les vecteurs $a$ et $a^k$ appartiennent au sous-espace $L$, ceci nous amènerait à conclure que le vecteur

$$y = a - \sum_{k=1}^{m} (a, a^k) a^k$$

appartient aussi à $L$. Or, le vecteur $y$ est orthogonal à tous les $a^s$ ($s = 1, \ldots, m$) (cf. § 17, (4)). Le vecteur

$$b = \frac{1}{|y|} y \tag{8}$$

appartiendrait aussi à $L$ et serait orthogonal à tous les $a^k$ ($k = 1, \ldots, m$), ce qui est impossible. Ceci prouve la proposition a) du théorème.

La complémentation du système orthonormal (4) à la base orthonormale (5) se fait à l'aide du théorème 1, § 17. Désignons par $L'$ le sous-espace de toutes les combinaisons linéaires $v = \sum_{k=m+1}^{n} \mu_k a^k$ des vecteurs du système (6). Chacun de ces vecteurs est visiblement orthogonal à tout vecteur $u \in L$ qui se représente par la somme $u = \sum_{k=1}^{m} \lambda_k a^k$. D'autre part, si $a \in R_n$ est un vecteur quelconque orthogonal à tous les vecteurs $u \in L$, en particulier à $a^1, \ldots, a^m$, alors son développement suivant la base (5) est de la forme

$$a = \sum_{k=1}^{n} (a, a^k) a^k = \sum_{k=m+1}^{n} (a, a^k) a^k,$$

c'est-à-dire que $a \in L'$. Ceci prouve la proposition b) du théorème.

Tout vecteur $u = \sum_{k=1}^{m} \lambda_k a^k \in L$ est orthogonal à tous les vecteurs $v = \sum_{k=m+1}^{n} \mu_k a^k \in L'$. Si l'on sait qu'un vecteur $a = \sum_{k=1}^{n} (a, a^k) \cdot a^k$ est orthogonal à tous les vecteurs de $L'$, notamment à $a^{m+1}, \ldots, a^n$, alors $a = \sum_{k=1}^{m} (a, a^k) a^k$, i.e. $a \in L$. Ceci prouve la proposition c).

Enfin, si $a \in R_n$ est un vecteur arbitraire, on peut le représenter d'une seule façon sous la forme d'une somme

$$a = \sum_{k=1}^{n} (a, a^k) a^k = u + v,$$

où

$$u = \sum_{k=1}^{m} (a, a^k) a^k \in L, \quad v = \sum_{k=m+1}^{n} (a, a^k) a^k \in L'.$$

Ceci achève la démonstration du théorème 1.

THÉORÈME 2. *Soit $L$ un sous-espace à $m$ dimensions de $R_n$. Alors le sous-espace $L' \subset R_n$ orthogonal à $L$ est à $n - m$ dimensions et de plus $L$ est à son tour orthogonal à $L'$.*

DÉMONSTRATION. Si $L$ est différent de $R_n$ et du sous-espace nul, ce théorème est de toute évidence exprimé dans le théorème 1. Si dans l'énoncé du théorème 1, on s'affranchit de tout ce qui concerne les bases (4), (5) et (6), on obtient le théorème 2.

Supposons que $L$ est un sous-espace nul. Tout vecteur $a \in R_n$ étant orthogonal à $0$, on a $L' = R_n$ et la dimension de $R_n$ est $n - 0 = n$. Inversement, le vecteur $0$ est orthogonal à tous les vecteurs $a \in R_n = L'$. Il n'existe pas d'autres vecteurs orthogonaux à tous les vecteurs de $R_n$, car tout vecteur différent de $0$ n'est pas orthogonal à lui-même. Nous avons montré que $L$ est orthogonal à $L'$.

Si $L = R_n$, on raisonne de façon analogue.

COROLLAIRE. *Soit donné un système de vecteurs*

$$x^1, \ldots, x^m, \tag{9}$$

*et soit $L'$ le sous-espace des vecteurs $v$ orthogonaux à ce système*

$$(v, x^k) = 0 \quad (k = 1, \ldots, m).$$

*Soit d'autre part un vecteur $a$ orthogonal à tous les vecteurs $v$, c'est-à-dire au sous-espace $L'$. Alors $a$ est une combinaison linéaire des vecteurs du système* (9)

$$a = \sum_{k=1}^{m} \lambda_k x^k.$$

DÉMONSTRATION. Considérons le sous-espace $L$ de toutes les combinaisons linéaires des vecteurs du système (9), autrement dit tout vecteur $u \in L$ est une combinaison linéaire

$$u = \sum_{k=1}^{m} \lambda_k x^k.$$

On dira aussi que le sous-espace $L$ est engendré par les vecteurs du système (9).

Les vecteurs $v \in L'$ étant orthogonaux à ceux du système (9), ils le sont visiblement à tout vecteur $u \in L$. Ceci indique que le sous-espace $L'$ est orthogonal à $L$. Donc, d'après le théorème 2, $L$ est aussi orthogonal à $L'$, autrement dit $L$ est composé de *tous* les vecteurs $u$ orthogonaux à $L'$. Par hypothèse, $a$ est un tel vecteur et par suite est une combinaison linéaire des vecteurs du système (9).

## § 21. Théorèmes de type Fredholm

Dans ce paragraphe on expose la théorie des équations linéaires parallèlement à la théorie développée au § 4.

Cette théorie ne fait pas appel aux déterminants : ceux-ci ne figurent pas explicitement dans les énoncés. Son avantage est d'avoir servi de base et d'analogue à de nombreuses généralisations en analyse mathématique, dont les premières et plus importantes sont dues à Fredholm [1]).

Nous considérons de nouveau un opérateur linéaire $A$ (voir § 15) :

$$y = Ax \quad (x \in R_n), \tag{1}$$

qui associe à tout vecteur $x \in R_n$ un vecteur $y \in R_n$ à l'aide des égalités

$$y_i = \sum_{s=1}^{n} a_{is} x_s \quad (i = 1, \ldots, n). \tag{2}$$

Ici

$$A = \| a_{is} \| = \begin{Vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{Vmatrix} \tag{3}$$

est une matrice carrée donnée. L'*adjoint* de $A$ est l'opérateur $A^*$ :

$$y = A^* x \quad (x \in R_n), \tag{1*}$$

défini par la matrice

$$A^* = \begin{Vmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ \cdot & \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{Vmatrix} \tag{3*}$$

adjointe de la matrice (2). On a

$$y_j = \sum_{l=1}^{n} a_{lj} x_l \quad (j = 1, \ldots, n), \tag{2*}$$

autrement dit la composante $y_j$ s'exprime en fonction des coordonnées du vecteur $x$ à l'aide de la $j$-ième ligne de la matrice $A^*$ ou de la $j$-ième colonne de la matrice $A$.

---

[1]) Ivar Fredholm, mathématicien suédois (1866-1927).

On a l'égalité

$$(Ax,\ z) = (x,\ A^*z), \quad \forall x,\ z \in R_n. \tag{4}$$

En effet

$$(Ax, z) = \sum_{i=1}^{n} y_i z_i = \sum_{i=1}^{n} \Big(\sum_{s=1}^{n} a_{is} x_s\Big) z_i = \sum_{s=1}^{n} \Big(\sum_{i=1}^{n} a_{is} z_i\Big) x_s = (x,\ A^*z).$$

L'égalité (4) est caractéristique de l'opérateur adjoint, car si pour un opérateur linéaire $B$ on avait

$$(Ax,\ z) = (x,\ Bz), \quad \forall x,\ z \in R_n, \tag{5}$$

alors nécessairement $B = A^*$. En effet ($B = \| b_{ik} \|$),

$$(Ax,\ z) = \sum_{i=1}^{n} \sum_{s=1}^{n} a_{is} x_s z_i,$$

$$(x,\ Bz) = \sum_{s=1}^{n} \sum_{i=1}^{n} b_{si} z_i x_s = \sum_{i=1}^{n} \sum_{s=1}^{n} b_{si} x_s z_i.$$

De (5) il résulte que

$$\sum_{i=1}^{n} \sum_{s=1}^{n} a_{is} x_s z_i = \sum_{i=1}^{n} \sum_{s=1}^{n} b_{si} x_s z_i, \ \forall x,\ z \in R_n, \tag{6}$$

d'où $a_{is} = b_{si}$ ($i,\ s = 1,\ \ldots,\ n$), ce qu'on vérifie immédiatement en faisant $x = (0,\ \ldots,\ 0,\ 1,\ 0,\ \ldots,\ 0)$ et $z = (0,\ \ldots,\ 0,\ 1,\ 0, \ldots$ $\ldots,\ 0)$ dans (6), l'unité se trouvant au $s$-ième rang dans $x$ et au $i$-ième dans $z$.

Donc, *on peut définir l'adjoint $A^*$ d'un opérateur linéaire $A$ comme un opérateur linéaire justiciable de l'égalité* (4).

Les égalités (1) et (1*) peuvent être traitées comme des équations: on donne un vecteur $y \in R_n$ et on cherche un vecteur $x \in R_n$ vérifiant (1) ou (1*).

Les équations homogènes correspondantes s'écrivent

$$Ax = 0 \tag{$1_0$}$$

ou

$$\sum_{s=1}^{n} a_{is} x_s = 0 \quad (i = 1,\ \ldots,\ n), \tag{$2_0$}$$

et

$$A^*z = 0 \tag{$1_0^*$}$$

ou

$$\sum_{l=1}^{n} a_{lj} z_l = 0 \quad (j = 1,\ \ldots,\ n). \tag{$2_0^*$}$$

Désignons par $L$ l'*image de l'espace $R_n$ par l'opérateur $A$*:

$$L = A\ (R_n),$$

et par $L'$ le *sous-espace des vecteurs* $z$ *vérifiant l'équation homogène adjointe* $(1_0^*)$.

Nous avons qualifié $L'$ de sous-espace, car il contient des vecteurs $z$, $z'$ avec leur combinaison linéaire $\alpha z + \beta z'$, où $\alpha$ et $\beta$ sont des nombres :

$$A^* (\alpha z + \beta z') = \alpha A^* z + \beta A^* z' = 0.$$

$L$ est aussi un sous-espace, car si $y$, $y' \in L$, alors il existe des vecteurs $x$, $x' \in R_n$, tels que $y = Ax$, $y' = Ax'$ et par suite

$$\alpha y + \beta y' = \alpha Ax + \beta Ax' = A (\alpha x + \beta x'),$$

c'est-à-dire que $\alpha y + \beta y' \in L$.

On a le lemme suivant (voir § 20, théorème 2).

Lemme. *Les sous-espaces* $L$ *et* $L'$ *sont mutuellement orthogonaux, autrement dit* $L'$ *est l'ensemble de tous les vecteurs* $z$ *orthogonaux à* $L$, *et* $L$ *celui de tous les vecteurs* $y$ *orthogonaux à* $L'$. *Si* $L$ *est de dimension* $k$, $L'$ *est de dimension* $n - k$.

Démonstration. Considérons l'égalité

$$(Ax, z) = (x, A^* z), \tag{7}$$

valable pour tous les $x$, $z \in R_n$. Supposons que $z$ est orthogonal à $L$ ; le premier membre de (7) est alors nul pour tous les $x \in R_n$, donc le second l'est aussi pour tous les $x \in R_n$ et en particulier pour $x = A^* z$ :

$$(A^* z, A^* z) = 0.$$

Par suite, $A^* z = 0$. Nous avons ainsi prouvé que si un vecteur $z$ est orthogonal à $L$, il vérifie alors l'équation $A^* z = 0$ (i.e. $z \in L'$).

Supposons inversement qu'un vecteur $z$ est solution de l'équation $A^* z = 0$. Ce vecteur annule le second membre de (7) quel que soit $x$, donc il annule le premier. Par conséquent, il est orthogonal à tous les vecteurs $Ax$, i.e. à tous les vecteurs $y \in L$ ou encore au sous-espace $L$.

Nous avons prouvé que $L'$ est l'ensemble de *tous* les vecteurs $z$ orthogonaux au sous-espace $L$. Mais alors le théorème 2 du § 20 nous dit qu'inversement $L$ est l'ensemble de *tous* les vecteurs $y$ orthogonaux à $L'$ et la somme des dimensions de $L$ et de $L'$ est $n$. Ceci achève la démonstration du lemme.

On a le théorème suivant.

Théorème 1. *Etant donné un vecteur* $y \in R_n$, *une condition nécessaire et suffisante pour que l'équation*

$$y = Ax \tag{1'}$$

*admette une solution est que le vecteur $y$ soit orthogonal à tous les vecteurs $z$, satisfaisant l'équation adjointe homogène*

$$A^*z = 0. \qquad (1_0^*)$$

*La solution $x$ de l'équation* (1), *si elle existe, se représente par la somme*

$$x = x^0 + u,$$

*où $x^0$ est une solution particulière de l'équation* (1), *et $u$ une solution quelconque de l'équation homogène*

$$Au = 0. \qquad (1_0)$$

*Toute somme $x^0 + u$ est solution de* (1').

DÉMONSTRATION. En vertu du lemme 1, si $L = A\ (R_n)$ et $L'$ est l'ensemble de tous les $z$ vérifiant l'équation $A^*z = 0$, alors $L$ et $L'$ sont des sous-espaces mutuellement orthogonaux. Si pour $y$ l'équation (1) admet une solution, alors $y \in L$ et tous les $z \in L'$ sont nécessairement orthogonaux à $y$. Si $y$ est orthogonal à tous les $z \in L'$, alors $y \in L$, autrement dit il existe un $x$ pour lequel $y = Ax$.

Supposons maintenant que l'équation (1') admet une solution $x^0$:

$$y = Ax^0.$$

Il est alors évident que la somme $x^0 + u$, où $Au = 0$, est aussi solution de l'équation (1'):

$$A\ (x^0 + u) = Ax^0 + Au = y + 0 = y.$$

Inversement, si $x$ est une solution arbitraire de l'équation (1') et $x^0$ une solution particulière déterminée, alors

$$y = Ax, \quad y = Ax^0,$$

et par suite

$$0 = Ax - Ax^0 = A\ (x - x^0) = Au,$$

où $u = x - x^0$, i.e. $x = x^0 + u$, où $u$ est solution de l'équation $Au = 0$.

REMARQUE. Mettons en évidence le lien qui relie le théorème 1 à la théorie de Kronecker-Capelli sur l'exemple de l'espace $R_2$. Supposons qu'un vecteur $y = (y_1, y_2)$ est orthogonal à toutes les solutions du système

$$\begin{cases} a_{11}z_1 + a_{21}z_2 = 0, \\ a_{12}z_1 + a_{22}z_2 = 0. \end{cases} \qquad (8)$$

Montrons alors que la matrice $A$ et la matrice élargie

$$B = \begin{Vmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \end{Vmatrix}$$

ont des rangs égaux. Si rang $A = 2$, il est évident que rang $B = 2$. Supposons que rang $A = 1$. On a toujours rang $B \geqslant$ rang $A = 1$. Il nous faut donc montrer que

$$\Delta_1 = \begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix} = 0, \quad \Delta_2 = \begin{vmatrix} a_{12} & y_1 \\ a_{22} & y_2 \end{vmatrix} = 0.$$

En effet, le vecteur $y$ étant orthogonal à toutes les solutions (non triviales) du système (8), il vient que $y_1 z_1 + y_2 z_2 = 0$. Donc, si l'on admet que $z_1 \neq 0$, on obtient

$$\Delta_1 = a_{11} y_2 - a_{21} y_1 = a_{11} y_2 + a_{21} \frac{y_2 z_2}{z_1} = \frac{y_2}{z_1} (a_{11} z_1 + a_{21} z_2) = 0,$$

$$\Delta_2 = a_{12} y_2 - a_{22} y_1 = a_{12} y_2 + a_{22} \frac{y_2 z_2}{z_1} = \frac{y_2}{z_1} (a_{12} z_1 + a_{22} z_2) = 0.$$

D'où il suit que rang $B$ = rang $A = 1$.

Inversement, supposons que $y = (y_1, y_2)$ est un vecteur tel que rang $B$ = rang $A$. Alors l'équation (1) admet une solution $(x_1, x_2)$. Montrons que $y$ est orthogonal à toutes les solutions $z = (z_1, z_2)$ du système (8). En effet,

$$y_1 z_1 + y_2 z_2 = (a_{11} x_1 + a_{12} x_2) z_1 + (a_{21} x_1 + a_{22} x_2) z_2 =$$
$$= (a_{11} z_1 + a_{21} z_2) x_1 + (a_{12} z_1 + a_{22} z_2) x_2 = 0 \cdot x_1 + 0 \cdot x_2 = 0.$$

Théorème 2. *Les équations homogènes*

$$Ax = 0 \qquad (1_0)$$

*et*

$$A^* x = 0 \qquad (1_0^*)$$

*admettent le même nombre de solutions linéairement indépendantes.*

*Si en particulier l'une de ces équations n'admet que la solution triviale* 0, *c'est-à-dire aucune solution indépendante, il en sera de même de l'autre.*

Remarque. L'équation (1) admet une solution unique dans le dernier cas.

Démonstration. Les matrices $A$ et $A^*$ sont de même rang $k$ et ont même déterminant $\Delta$.

Si $k = n$, alors $\Delta \neq 0$ et les équations $(1_0)$ et $(1_0^*)$ n'admettent que la solution triviale 0. Dans ce cas, le théorème 1 nous dit que l'équation (1) admet une solution unique quel que soit $y \in R_n$.

Supposons maintenant que $1 \leqslant k < n$. Renumérotons les variables $x_n$ de telle sorte que

$$\begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} \neq 0.$$

Ecrivons les $k$ premières équations ($1_0$) sous la forme

$$\begin{cases} a_{11}x_1 + \ldots + a_{1k}x_k = -a_{1,\,k+1}x_{k+1} - \ldots - a_{1n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{k1}x_1 + \ldots + a_{kk}x_k = -a_{k,\,k+1}x_{k+1} - \ldots - a_{kn}x_n. \end{cases} \quad (9)$$

Considérons les $n-k$ vecteurs

$$\begin{cases} x^1 = (x_1^1, \ldots, x_k^1, 1, 0, \ldots, 0), \\ x^2 = (x_1^2, \ldots, x_k^2, 0, 1, 0, \ldots, 0), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^{n-k} = (x_1^{n-k}, \ldots, x_k^{n-k}, 0, \ldots, 0, 1). \end{cases} \quad (10)$$

Pour obtenir le premier vecteur on porte

$$x_{k+1} = 1, \quad x_{k+2} = 0, \ldots, x_n = 0$$

dans le système (9) et on le résout par rapport à $x_1, \ldots, x_k$. Désignons par $x_1^1, \ldots, x_k^1$ les solutions uniques obtenues. Pour obtenir le deuxième vecteur, on porte

$$x_{k+1} = 0, \quad x_{k+2} = 1, \quad x_{k+3} = 0, \ldots, x_n = 0$$

dans (9) et l'on trouve les solutions $x_1^2, \ldots, x_k^2$, et ainsi de suite. Les vecteurs (10) sont doués des propriétés suivantes:

1) Le système de vecteurs (10) est linéairement indépendant, car le rang de sa matrice est égal au nombre des vecteurs $\mu = n - k$.

2) Chaque vecteur du système (10) est solution de l'une quelconque des équations ($1_0$) ou $Ax = 0$.

3) Les solutions de l'équation $Ax = 0$ sont toutes de la forme

$$\lambda_1 x^1 + \ldots + \lambda_{n-k} x^{n-k},$$

où $\lambda_1, \ldots, \lambda_{n-k}$ sont des nombres arbitraires.

On exprime ces trois propositions en disant que l'équation ($1_0$) admet $n-k$ solutions linéairement indépendantes.

Par des raisonnements analogues et en tenant compte de ce que rang $A$ = rang $A^*$, on démontre que l'équation $A^*x = 0$ admet aussi $n - k$ solutions linéairement indépendantes. C.q.f.d.

THÉORÈME 3. *Si l'une des équations homogènes* ($1_0$) *ou* ($1_0^*$) *admet $k$ solutions linéairement indépendantes, alors l'autre en admet aussi $k$; les images $L = A(R_n)$ et $L^* = A^*(R_n)$ de $R_n$ par les opérateurs $A$ et $A^*$ sont des sous-espaces à $n - k$ dimensions.*

DÉMONSTRATION. La première proposition fait l'objet du théorème 2, la seconde, du lemme 1 qui dit que le sous-espace $L$ est à $n - k$ dimensions, où $k$ est la dimension du sous-espace $L'$ des vecteurs $z$, solutions de l'équation $A^*z = 0$. De façon analogue, le sous-espace $L^*$ est à $n - k$ dimensions, où $k$ est le nombre de dimensions du sous-espace des vecteurs $u$ vérifiant l'équation $Au = 0$.

## § 22. Opérateur autoadjoint. Forme quadratique

L'opérateur linéaire

$$y_k = \sum_{l=1}^{n} a_{kl}x_l \qquad (k = 1, \ldots, n) \tag{1}$$

ou de façon plus concise

$$y = Ax \quad (x \in R_n, \ y \in R_n) \tag{2}$$

est dit *autoadjoint* ou *hermitien* s'il est égal à son adjoint, c'est-à-dire si

$$Ax = A^*x, \quad \forall\, x \in R_n, \tag{3}$$

en d'autres termes, si la matrice $A$ est symétrique:

$$a_{kl} = a_{lk} \quad (k,\ l = 1, \ldots, n) \tag{4}$$

(voir (3) et (3*), § 21).

L'opérateur $A$ est justiciable de l'égalité caractéristique

$$(x,\ Az) = (Ax,\ z), \quad \forall x,\ z \in R_n$$

(voir § 21, (4)). Il est évident que

$$(x,\ Ax) = \sum_{k=1}^{n} x_k \sum_{l=1}^{n} a_{kl}x_l = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl}x_kx_l \quad (a_{kl} = a_{lk}). \tag{4'}$$

L'expression du second membre de (4') s'appelle *forme quadratique d'ordre n*. C'est une fonction continue du vecteur $x$ ou ce qui revient au même des variables $x_1, \ldots, x_n$.

On étudiera cette fonction sur l'ensemble $S$ de valeurs des vecteurs $x$ de norme un ($|x| = 1$). L'ensemble $S$ est une sphère de $R_n$, de rayon 1 et de centre 0. L'ensemble $S$ est borné. Il est également fermé [1]), c'est-à-dire que si les points d'une suite $\{x^\nu\}$ ($\nu = 1, 2, \ldots$) appartiennent à $S$ ($|x^\nu| = 1$, $\nu = 1, 2, \ldots$) et si cette suite converge vers un point $x^0 \in R_n$ ($x^\nu \to x^0$, $\nu \to \infty$), alors $x^0 \in S$, i.e. $|x^0| = 1$, car $|1 - |x^0|| = ||x^\nu| - |x^0|| \leqslant |x^\nu - x^0| \to 0$, d'où $|x^0| = 1$.

Calculons le maximum de la forme quadratique (4') sur la sphère $S$. La forme (4') étant une fonction continue sur un ensemble borné fermé, son maximum est réalisé pour un vecteur unité $x^1$. Soit $\lambda_1$ ce maximum:

$$\lambda_1 = (Ax^1,\ x^1) \geqslant (Ax,\ x), \quad \forall x: |x| = 1. \tag{5}$$

Considérons le sous-espace $L'$ orthogonal au vecteur $x^1$, c'est-à-dire l'ensemble de tous les vecteurs $v$ orthogonaux à $x^1$. Soit $v^0$ un vecteur

[1]) Voir chap. 8, § 12.

unité de $L'$. Le vecteur

$$\cos\alpha\cdot\boldsymbol{x}^1+\sin\alpha\cdot\boldsymbol{v}^0$$

dépend de $\alpha$ et sa norme est égale à un:

$$|\cos\alpha\cdot\boldsymbol{x}^1+\sin\alpha\cdot\boldsymbol{v}^0|=$$
$$=(\cos\alpha\cdot\boldsymbol{x}^1+\sin\alpha\cdot\boldsymbol{v}^0,\ \cos\alpha\cdot\boldsymbol{x}^1+\sin\alpha\cdot\boldsymbol{v}^0)^{1/2}$$
$$=(\cos^2\alpha+\sin^2\alpha)^{1/2}=1.$$

Pour $\alpha=0$, ce vecteur se confond avec $\boldsymbol{x}^1$. Mais alors la fonction

$$\psi(\alpha)=(A(\cos\alpha\cdot\boldsymbol{x}^1+\sin\alpha\cdot\boldsymbol{v}^0),\ \cos\alpha\cdot\boldsymbol{x}^1+\sin\alpha\cdot\boldsymbol{v}^0)$$

atteint son maximum au point $\alpha=0$ ($\psi(0)=(A\boldsymbol{x}^1,\ \boldsymbol{x}^1)$) et en vertu de la condition nécessaire d'extrémum

$$\psi'(0)=0.$$

Calculons cette dérivée. On a

$$\psi(\alpha)=\cos^2\alpha\,(A\boldsymbol{x}^1,\ \boldsymbol{x}^1)+\sin 2\alpha\,(A\boldsymbol{x}^1,\ \boldsymbol{v}^0)+\sin^2\alpha\,(A\boldsymbol{v}^0,\ \boldsymbol{v}^0).$$

Donc

$$\psi'(\alpha)=-\sin 2\alpha\,(A\boldsymbol{x}^1,\ \boldsymbol{x}^1)+2\cos 2\alpha\,(A\boldsymbol{x}^1,\ \boldsymbol{v}^0)+$$
$$+\sin 2\alpha\,(A\boldsymbol{v}^0,\ \boldsymbol{v}^0)$$

et

$$\psi'(0)=2\,(A\boldsymbol{x}^1,\ \boldsymbol{v}^0)=0.$$

On voit que le vecteur $A\boldsymbol{x}^1$ est orthogonal à tout vecteur unité $\boldsymbol{v}^0\in L'$, donc à tout vecteur $\boldsymbol{v}\in L'$. Alors $A\boldsymbol{x}^1$ diffère de $\boldsymbol{x}^1$ d'un facteur multiplicatif (voir corollaire 1, fin du § 20), autrement dit

$$A\boldsymbol{x}^1=\lambda\boldsymbol{x}^1,$$

où $\lambda$ est un nombre quelconque.

De la première relation (l'égalité) de (5) et puisque $|\boldsymbol{x}^1|=1$, il vient

$$\lambda_1=(\lambda\boldsymbol{x}^1,\ \boldsymbol{x}^1)=\lambda.$$

Nous avons ainsi démontré que le maximum de la forme quadratique (4′) sur la sphère $S$ est atteint en un point $\boldsymbol{x}^1$:

$$\max_{|\boldsymbol{x}|=1}\ (A\boldsymbol{x},\ \boldsymbol{x})=(A\boldsymbol{x}^1,\ \boldsymbol{x}^1)=\lambda_1.$$

Ceci étant,

$$A\boldsymbol{x}^1=\lambda_1\boldsymbol{x}^1,\qquad |\boldsymbol{x}^1|=1.$$

On voit donc que l'image du vecteur non trivial $\boldsymbol{x}^1$ par l'opérateur $A$ est le vecteur colinéaire $\lambda_1\boldsymbol{x}^1$.

Un tel vecteur s'appelle *vecteur propre de l'opérateur* $A$ et le nombre $\lambda_1$ *valeur propre associée au vecteur* $x^1$.

On considérera maintenant l'opérateur $A$ sur un sous-espace $R^1$ défini comme l'ensemble de tous les vecteurs $x \in R_n$ orthogonaux à $x^1$ (ce sous-espace a été désigné plus haut par $L'$). Le sous-espace $R^1$ est à $n-1$ dimensions. Il contient des bases orthonormales de $n-1$ vecteurs. On se propose d'en trouver une, qui, on le verra, est liée de façon intrinsèque à l'opérateur $A$.

Il importe de souligner que l'image $A(R^1)$ de $R^1$ par l'opérateur $A$ appartient à $R^1$, car si $(x, x^1) = 0$, alors

$$(Ax, x^1) = (x, Ax^1) = (x, \lambda_1 x^1) = \lambda_1 (x, x^1) = 0,$$

c'est-à-dire que $Ax \in R^1$.

L'hermicité de l'opérateur $A$ est préservée de façon triviale sur $R^1$, car l'égalité

$$(Ax, y) = (x, Ay),$$

valable pour tous les $x$, $y \in R_n$, l'est également pour $x$, $y \in R^1$.

On peut donc étudier maintenant un opérateur linéaire hermitien $A$ sur le sous-espace vectoriel $R^1$ à $n-1$ dimensions. On peut appliquer à cet opérateur les raisonnements précédents et exhiber dans $R^1$ un vecteur unité $x^2$ tel que

$$\max_{\substack{|x|=1 \\ (x, x^1)=0}} (Ax, x) = (Ax^2, x^2)_{\substack{|x^2|=1 \\ (x^2, x^1)=0}} = \lambda_2 \leqslant \lambda_1.$$

En effet, la sphère unité $S^1$ de $R^1$ se définit de toute évidence comme l'ensemble de tous les vecteurs unités orthogonaux à $x^1$. En outre,

$$Ax^2 = \lambda_2 x^2.$$

Nous avons trouvé un deuxième vecteur propre de l'opérateur $A$ et la valeur propre $\lambda_2$ associée à $x^2$. De toute évidence, celle-ci n'est pas supérieure à $\lambda_1$ (si l'on réduit le domaine, le maximum ne peut que diminuer). Ceci étant, $(x^1, x^2) = 0$.

On peut par analogie introduire un sous-espace $R^2$ à $n-2$ dimensions, orthogonal aux vecteurs $x^1$ et $x^2$, montrer que $A$ laisse $R^2$ invariant et exhiber un vecteur unité $x^3$ orthogonal à $x^1$ et $x^2$, tel que

$$\max_{\substack{|x|=1 \\ (x, x^1)=0,\ (x, x^2)=0}} (Ax, x) = (Ax^3, x^3) = \lambda_3$$

et

$$Ax^3 = \lambda_3 x^3 \quad (\lambda_3 \leqslant \lambda_2 \leqslant \lambda_1).$$

En poursuivant cette procédure jusqu'à $x^n$, on obtient un système orthonormal de vecteurs

$$x^1, x^2, \ldots, x^n \tag{6}$$

et un système de nombres réels

$$\lambda_1, \lambda_2, \ldots, \lambda_n, \tag{7}$$

tels que

$$\begin{cases} Ax^k = \lambda_k x^k \quad (k=1, \ldots, n), \\ \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n. \end{cases} \tag{8}$$

Remarque 1. Si l'on considère un opérateur hermitien $A$ sur un espace complexe $R_n$, on démontre que ses valeurs propres sont réelles. En effet, soient $x$ un vecteur propre de $A$ et $\lambda$ la valeur propre associée à $x$, c'est-à-dire que $Ax = \lambda x$, $x \neq 0$. L'opérateur $A$ étant hermitien on a $(Ax, x) = (x, Ax)$ ou $(\lambda x, x) = (x, \lambda x)$. De là, en utilisant les propriétés du produit scalaire dans l'espace complexe $R_n$ (cf. § 6) on déduit que $\lambda (x, x) = \overline{\lambda} (x, x)$. Comme $(x, x) \neq 0$, il vient $\lambda = \overline{\lambda}$, c'est-à-dire que $\lambda$ est réel.

Par ailleurs, les vecteurs propres de l'opérateur hermitien auxquels sont associées des valeurs propres distinctes sont orthogonaux. En effet:

$$Ax^1 = \lambda_1 x^1, \quad Ax^2 = \lambda_2 x^2 \quad (\lambda_2 \neq \lambda_1),$$
$$(Ax^1, x^2) = (\lambda_1 x^1, x^2) = \lambda_1 (x^1, x^2),$$
$$(x^1, Ax^2) = (x^1, \lambda_2 x^2) = \overline{\lambda}_2 (x^1, x^2) = \lambda_2 (x^1, x^2)$$

et

$$0 (\lambda_1 - \lambda_2)(x^1, x^2).$$

Comme $\lambda_1 \neq \lambda_2$, il s'ensuit que $(x^1, x^2) = 0$, ce qui exprime que les vecteurs $x^1$ et $x^2$ sont orthogonaux.

Nous avons obtenu un *système complet de vecteurs propres de l'opérateur $A$* ainsi que les valeurs propres correspondantes. Le système orthonormal (6) est une base de $R_n$, car il est composé de $n$ vecteurs appartenant à $R_n$ (voir § 17). Donc, tout vecteur $x \in R_n$ peut être développé suivant ce système

$$x = \sum_{k=1}^{n} (x, x^k) x^k. \tag{9}$$

L'opérateur hermitien $A$ peut alors s'écrire

$$Ax = A\left(\sum_{k=1}^{n} (x, x^k) x^k\right) = \sum_{k=1}^{n} (x, x^k) Ax^k = \sum_{k=1}^{n} \lambda_k (x, x^k) x^k. \tag{10}$$

Nous avons démontré le

Théorème 1. *A tout opérateur hermitien $A$ de $R_n$ est associé un système orthogonal de vecteurs $x^1, \ldots, x^n$ (une base de $R_n$) et un système de nombres réels $\lambda_1, \ldots, \lambda_n$, tels que $Ax$ se représente par une somme* (10) *quel que soit $x \in R_n$.*

La forme quadratique (4′) s'écrit alors:

$$(x, Ax) = \left(\sum_{k=1}^{n}(x, x^k)\,x^k, \ \sum_{s=1}^{n}\lambda_s (x, x^s)\,x^s\right) = \sum_{s=1}^{n}\lambda_s (x, x^s)^2. \qquad (4'')$$

Pour appliquer les résultats obtenus à la forme quadratique (4′), on peut définir l'opérateur linéaire

$$y = Ax$$

à l'aide des égalités

$$y_i = \sum_{j=1}^{n} a_{ij}x_j \quad (i = 1, \ldots, n).$$

Comme $a_{kl} = a_{lk}$, cet opérateur est hermitien et on peut lui appliquer le théorème 1. Dans le langage des formes quadratiques, le théorème 1 s'énonce comme suit:

THÉORÈME 2. *Soit donnée une forme quadratique* (4′) *dans un espace $R_n$ à $n$ dimensions, rapporté à un système de coordonnées* $(x_1, \ldots, x_n)$ *de vecteurs unités* $i^1, \ldots, i^n$ $(x = \sum_{k=1}^{n} x_k i^k)$. *Il existe un système de coordonnées rectangulaires* $(\xi_1, \ldots, \xi_n)$ *de vecteurs unités* $x^1, \ldots, x^n$, *formant une base orthogonale* $(x = \sum_{k=1}^{n} \xi_k x^k)$ *et un système de nombres réels* $\lambda_1, \ldots, \lambda_n$, *tels que la forme quadratique* (4′) *est, dans ce système, la somme des carrés des coordonnées* $\xi_s$ *du vecteur* $x$, *multipliés respectivement par les nombres* $\lambda_s$:

$$\sum_{k=1}^{n}\sum_{l=1}^{n} a_{kl}x_k x_l = \sum_{s=1}^{n}\lambda_s \xi_s^2. \qquad (4''')$$

On passe du premier membre de (4‴) au second si l'on connaît les développements des vecteurs $x^1, \ldots, x^n$ suivant les vecteurs unités $i^1, \ldots, i^n$. Soit

$$x^j = \sum_{s=1}^{n}\beta_{js} i^s$$

(voir § 17, (7), où il faut remplacer $a_{js}$ et $a^s$ respectivement par $\beta_{js}$ et $i^s$). Comme $i^1, \ldots, i^n$ et $x^1, \ldots, x^n$ sont des bases orthonormales de $R_n$, la matrice

$$\Lambda = \| \beta_{js} \|$$

est orthogonale. On admet qu'elle est connue. Développons $x$ suivant les deux bases:

$$x = \sum_{s=1}^{n} x_s i^s = \sum_{j=1}^{n}\xi_j x^j.$$

Donc

$$\sum_{j=1}^{n} \xi_j x^j = \sum_{j=1}^{n} \xi_j \sum_{s=1}^{n} \beta_{js} i^s = \sum_{s=1}^{n} \Big( \sum_{j=1}^{n} \beta_{js}\xi_j \Big) i^s$$

et, puisque le système $i^1, \ldots, i^n$ est linéairement indépendant,

$$x_s = \sum_{j=1}^{n} \beta_{js}\xi_j \quad (s = 1, \ldots, n). \tag{11}$$

On passe donc des coordonnées $\xi_1, \ldots, \xi_n$ aux coordonnées $x_1, \ldots, x_n$ à l'aide de la matrice $\Lambda^*$ adjointe de $\Lambda$ (c'est-à-dire à l'aide des lignes de $\Lambda^*$ ou des colonnes de $\Lambda$).

Si l'on porte l'expression (11) de $x_s$ dans le premier membre de (4‴), on doit obtenir le second. Écrivons cette égalité:

$$\sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} \sum_{j=1}^{n} \beta_{jk}\xi_j \sum_{i=1}^{n} \beta_{il}\xi_i = \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{k=1}^{n} \beta_{jk} \Big( \sum_{l=1}^{n} a_{kl}\beta_{il} \Big) \xi_j\xi_i =$$

$$= \sum_{s=1}^{n} \lambda_s \xi_s^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} \delta_{ji}\lambda_i \xi_j \xi_i,$$

où $\delta_{ji} = \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases}$ est le symbole de Kronecker.

L'identification des coefficients en $\xi_j\xi_i$ nous donne les égalités

$$\sum_{k=1}^{n} \beta_{jk} \Big( \sum_{l=1}^{n} a_{kl}\beta_{il} \Big) = \delta_{ji}\lambda_i \quad (i, j = 1, \ldots, n)$$

que l'on peut interpréter comme suit (voir § 15, (6)). Pour la matrice

$$A = \| a_{kl} \| \quad (a_{kl} = a_{lk})$$

de l'opérateur hermitien il existe une matrice orthogonale

$$\Lambda = \| \beta_{js} \|$$

telle que

$$\Lambda \cdot A \cdot \Lambda^{-1} = \mathfrak{A}, \tag{12}$$

où $\mathfrak{A}$ est une matrice diagonale

$$\mathfrak{A} = \begin{Vmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \lambda_n \end{Vmatrix} \tag{13}$$

(les $\lambda_i$ sont réels) dite *canonique*.

Signalons que pour la matrice orthogonale $\Lambda$ on a

$$\Lambda^{-1} = \Lambda^*.$$

Comme les déterminants des matrices orthogonales sont égaux à $\pm 1$, il s'ensuit de (12) que

$$\prod_{j=1}^{n} \lambda_j = |\mathfrak{A}| = |\Lambda| \cdot |A| \cdot |\Lambda^{-1}| = |A|. \tag{14}$$

Nous avons prouvé le théorème suivant.

THEOREME 3. *Si le déterminant $|A|$ de la matrice hermitienne $A$ n'est pas nul ($|A| \neq 0$), ses valeurs propres $\lambda_1, \ldots, \lambda_n$ sont toutes non nulles ($\lambda_j \neq 0$, $j = 1, \ldots, n$).*

Du théorème 2 il résulte que

1) Si $\lambda_1 \geqslant \ldots \geqslant \lambda_n > 0$, alors la forme quadratique est positive quel que soit $\xi \neq 0$, donc quel que soit $x \neq 0$. On dit qu'elle est *strictement positive.*

2) Si $0 > \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, alors la forme est négative quel que soit $\xi \neq 0$, donc, quel que soit $x \neq 0$. On dit qu'elle est *strictement négative.*

3) Si $\lambda_1 \geqslant \ldots \geqslant \lambda_n$ et $\lambda_n = 0$, alors la forme est positive. Il existe une direction (l'axe $\xi_n$) le long de laquelle elle s'annule. On dit que cette forme est *positive.*

4) Si $\lambda_1 = 0 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$, la forme est *négative.*

5) Si $\lambda_1 > 0$ et $\lambda_n < 0$, la forme n'est pas *définie.* Si l'on exclut le point $\mathbf{0}$, cette forme est positive le long de l'axe des $\xi_1$ et négative le long de l'axe des $\xi_n$.

Il existe un théorème, le *théorème de Sylvester* [1]), qui d'après la forme et le signe de certains déterminants d'une matrice $\|A\|$ permet de reconnaître si ses valeurs propres sont toutes positives ou toutes négatives ou certaines positives et les autres négatives.

Considérons les principaux mineurs de la forme quadratique $(Ax, x)$:

$$\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \ldots,\ \Delta_n = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}.$$

Le théorème de Sylvester, que nous ne démontrerons pas, dit que

1. Si $\Delta_1 > 0$, $\Delta_2 > 0$, $\ldots$, $\Delta_n > 0$, la forme est strictement positive (cas 1).

2. Si $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, $\ldots$, $(-1)^n \Delta_n > 0$, la forme est strictement négative (cas 2).

3. Si $\Delta_1 \geqslant 0$, $\Delta_2 \geqslant 0$, $\ldots$, $\Delta_n \geqslant 0$ ou

$$\Delta_1 \leqslant 0,\ \Delta_2 \geqslant 0,\ \ldots,\ (-1)^n \Delta_n \geqslant 0$$

---

[1]) James Joseph Sylvester, mathématicien anglais (1814-1897).

et il existe un $j$ pour lequel $\Delta_j = 0$, alors à fortiori la forme n'est pas strictement positive ou négative.

4. Dans tous les autres cas, la forme quadratique n'est pas définie.

## § 23. Forme quadratique dans un espace à deux dimensions

Pour $n = 2$, la forme quadratique s'écrit

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_2x_1 + a_{22}x_2^2 = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 \quad (1)$$

puisque $a_{12} = a_{21}$.

Pour réduire la forme (1) à une somme des carrés des coordonnées du vecteur $(\xi_1, \xi_2)$ dans une base $(x^1, x^2)$, il faut (voir § 22) trouver les vecteurs unités de la base $(x^1, x^2)$, c'est-à-dire les vecteurs propres de l'opérateur hermitien défini par la matrice symétrique

$$A = \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}.$$

Indiquons une méthode de recherche des valeurs propres et des vecteurs propres de l'opérateur $A$, différente de la méthode du § 22.

Si $\lambda_0$ est une valeur propre de l'opérateur $A$ et $x^0 = (x_1^{(0)}, x_2^{(0)}) \neq$ $\neq 0$ le vecteur propre correspondant, alors

$$Ax^0 = \lambda_0 x^0.$$

Cette équation s'écrit:

$$\begin{cases} (a_{11} - \lambda_0)\, x_1^{(0)} + a_{12}x_2^{(0)} = 0, \\ a_{21}x_1^{(0)} + (a_{22} - \lambda_0)\, x_2^{(0)} = 0, \end{cases} \quad (2)$$

ou sous la forme opératorielle

$$(A - \lambda_0 E)\, x^0 = 0, \quad (2')$$

où $E$ est l'opérateur identique.

Donc, le système homogène (2) admet une solution non triviale $x^0$ si seulement le déterminant du système (2) ou (2′) est nul:

$$\begin{vmatrix} a_{11} - \lambda_0 & a_{12} \\ a_{12} & a_{22} - \lambda_0 \end{vmatrix} = | A - \lambda_0 E | = 0.$$

Par suite. la valeur propre $\lambda_0$ est racine de l'équation

$$| A - \lambda_0 E | = 0, \quad (3)$$

appelée *équation caractéristique* de l'opérateur $A$ (ou de la forme quadratique $(Ax, x)$).

La réciproque est vraie. Si $\lambda_0$ est racine de l'équation (3), alors la solution non triviale du système

$$(A - \lambda_0 E)x = 0 \tag{4}$$

sera vecteur propre de l'opérateur hermitien $A$.

Donc, les valeurs propres de l'opérateur $A$ sont ici les racines de l'équation du second degré (3)

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}^2 = 0$$

ou encore

$$\lambda^2 - (a_{11} \quad a_{22})\lambda + a_{11}a_{22} - a_{12}^2 = 0.$$

Les racines de cette équation sont:

$$\begin{cases} \lambda_1 = \frac{1}{2}[a_{11} + a_{22} + \sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}], \\ \lambda_2 = \frac{1}{2}[a_{11} + a_{22} - \sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}]. \end{cases} \tag{5}$$

On voit que $\lambda_1 \geqslant \lambda_2$ et que $\lambda_1 = \lambda_2$ si $a_{12} = 0$, $a_{11} = a_{22}$. On admettra pour fixer les idées que $a_{11} \geqslant a_{22}$ (sinon on permute $x_1$ et $x_2$). Alors $\lambda_1 \geqslant a_{11}$ $(\lambda_1 - a_{11} = \frac{1}{2}[\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2} - (a_{11} - - a_{22})] \geqslant 0)$.

De (5) il résulte que les *valeurs propres de l'opérateur (hermitien) $A$ sont réelles.*

Trouvons maintenant les vecteurs unités propres associés à $\lambda_1$ et $\lambda_2$ en tant que solutions du système (4). Comme $|A - \lambda_1 E| = 0$, on a

$$\text{rang}\,(A - \lambda_1 E) \leqslant 1.$$

Si $\lambda_1 = \lambda_2$, la matrice $A - \lambda_1 E$ est entièrement composée de zéros $(\lambda_1 = \lambda_2 = a_{11} = a_{22}, a_{12} = 0)$ et son rang est nul. Dans ce cas, la forme quadratique est déjà réduite à une somme de carrés $(a_{12} = a_{21} = 0)$. Le système (4) est vérifié par tout vecteur $(x_1, x_2)$. Donc, pour vecteurs propres on peut prendre les vecteurs unités du système de coordonnées $\boldsymbol{x}^1 = \boldsymbol{i} = (1, 0)$, $\boldsymbol{x}^2 = \boldsymbol{j} = (0, 1)$. Dans tout autre système orthonormal $(\boldsymbol{x}^1, \boldsymbol{x}^2)$ la forme quadratique se réduit aussi à une somme de carrés.

Si $\lambda_1 > \lambda_2$, alors ou bien $a_{12} \neq 0$, ou bien $a_{12} = 0$, $a_{11} \neq a_{22}$. On peut ne pas traiter le deuxième cas, puisque la forme (1) a déjà été réduite à une somme de carrés. Supposons donc que $a_{12} \neq 0$. Alors

$$\text{rang}\,(A - \lambda_1 E) = 1.$$

Il suffit par conséquent de considérer une des équations du système (4):

$$(a_{11} - \lambda_1)x_1 + a_{12}x_2 = 0.$$

D'où (puisque $a_{12} \neq 0$)

$$x_2 = [(-a_{11} + \lambda_1)/a_{12}]\, x_1.$$

Le vecteur

$$y^1 = \left(x_1,\ \frac{\lambda_1 - a_{11}}{a_{12}} x_1\right)$$

est solution du système (4). En le normant on obtient le vecteur propre

$$x^1 = (x_1^{(1)}, x_2^{(1)}) = \frac{y^1}{|y^1|} =$$

$$= \left( \frac{\pm 1}{\sqrt{1 + \left(\frac{\lambda_1 - a_{11}}{a_{12}}\right)^2}},\ \frac{\pm(\lambda_1 - a_{11})}{a_{12}\sqrt{1 + \left(\frac{\lambda_1 - a_{11}}{a_{12}}\right)^2}} \right).$$

Des transformations élémentaires nous donnent :

$$\left\{ \begin{aligned} x_1^{(1)} &= \pm \frac{1}{\sqrt{2}} \sqrt{1 + (a_{11} - a_{22})/\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}}, \\ x_2^{(1)} &= \pm \frac{\operatorname{sgn} a_{12}}{\sqrt{2}} \sqrt{1 - (a_{11} - a_{22})/\sqrt{4a_{12}^2 + (a_{11} - a_{22})^2}}. \end{aligned} \right. \tag{6}$$

Dans la suite il suffit de prendre le signe + dans les formules (6).

On détermine de façon analogue le vecteur propre $x^2$ associé à la valeur propre $\lambda_2$. Il se trouve que

$$x^2 = (x_1^{(2)},\ x_2^{(2)}) = (-x_2^{(1)},\ x_1^{(1)}).$$

Composons maintenant la matrice de l'opérateur (de la transformation orthogonale) $\Lambda$ de passage de la base $(i, j)$ à la base $(x^1, x^2)$:

$$\Lambda = \begin{Vmatrix} x_1^{(1)} & x_2^{(1)} \\ -x_2^{(1)} & x_1^{(1)} \end{Vmatrix}$$

(les éléments des lignes sont les coordonnées des images des vecteurs unités $i$ et $j$ par $\Lambda$, c'est-à-dire que $x^1 = x_1^{(1)} i + x_2^{(1)} j$, $x^2 = -x_2^{(1)} i + x_1^{(1)} j$). Les coordonnées du vecteur $(x_1, x_2)$ dans le système $(i, j)$ sont reliées aux coordonnées $(\xi_1, \xi_2)$ de ce vecteur dans le système $(x^1, x^2)$ à l'aide des colonnes de la matrice $\Lambda$:

$$\left\{ \begin{aligned} x_1 &= x_1^{(1)} \xi_1 - x_2^{(1)} \xi_2, \\ x_2 &= x_2^{(1)} \xi_1 + x_1^{(1)} \xi_2. \end{aligned} \right. \tag{7}$$

En portant ces valeurs dans la forme quadratique (1) et en tenant compte des formules (5) et (6), on obtient

$$a_{11} x_1^2 + 2a_{12} x_1 x_2 + a_{22} x_2^2 = \lambda_1 \xi_1^2 + \lambda_2 \xi_2^2. \tag{8}$$

Le second membre de cette égalité s'appelle *forme canonique de la forme quadratique* (1).

On dira que la forme quadratique est de *type elliptique* si $\lambda_1$ et $\lambda_2$ sont de même signe, de *type hyperbolique* s'ils sont de signe contraire et de *type parabolique* si l'un d'eux est nul.

De (5) on voit que $\lambda_1\lambda_2 = a_{11}a_{22} - a_{12}^2$. Donc, le type de la forme (1) peut être défini d'après le signe de l'expression $a_{11}a_{22} - a_{12}^2$.

La forme quadratique est *elliptique*, *hyperbolique* ou *parabolique* selon que l'expression $a_{11}a_{22} - a_{12}^2$ est positive, négative ou nulle.

Exemple. Réduire à sa forme canonique

$$x_1^2 - \sqrt{3}\, x_1 x_2 + 2x_2^2.$$

Ici $a_{11} = 1$, $a_{12} = -\frac{\sqrt{3}}{2}$, $a_{22} = 2$. Comme $a_{11}a_{22} - a_{12}^2 = 2 - -\frac{3}{4} = \frac{5}{4} > 0$, la forme est elliptique. Trouvons les vecteurs et les valeurs propres à l'aide des formules (5) et (6):

$$x_1^{(1)} = \frac{1}{2}, \quad x_2^{(1)} = -\frac{\sqrt{3}}{2}, \quad x^1 = \left(\frac{1}{2}, \; -\frac{\sqrt{3}}{2}\right), \quad \lambda_1 = \frac{5}{2}.$$

D'autre part,

$$x^2 = \left(\frac{\sqrt{3}}{2}, \; \frac{1}{2}\right), \quad \lambda_2 = \frac{1}{2}.$$

Dans le système $(x^1, x^2)$, la forme quadratique s'écrit

$$\frac{5}{2}\xi_1^2 + \frac{1}{2}\xi_2^2.$$

Comme $x_1^{(1)} = \frac{1}{2} = \cos\left(-\frac{\pi}{3}\right)$, $\quad x_2^{(1)} = -\frac{\sqrt{3}}{2} = \sin\left(-\frac{\pi}{3}\right)$, la transformation à laquelle est associée la matrice

$$\Lambda = \begin{Vmatrix} x_1^{(1)} & -x_2^{(1)} \\ x_2^{(1)} & x_1^{(1)} \end{Vmatrix} \quad (x^1 = (x_1^{(1)}, \; x_2^{(1)}), \quad x^2 = (-x_2^{(1)}, \; x_1^{(1)}))$$

est une rotation du système de coordonnées $(Ox_1, Ox_2)$ d'un angle $\alpha = \pi/3$ autour de l'origine dans le sens des aiguilles d'une montre (voir exemple, fin du § 16).

## § 24. Coniques

Dans un plan rapporté à un système de coordonnées rectangulaires $(x, y)$, soit donnée une courbe définie implicitement par l'équation du second degré

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \tag{1}$$

où $A$, $B$, $C$, $D$, $E$, $F$ sont des nombres réels donnés, les nombres $A$, $B$, $C$ n'étant pas tous nuls. Cette courbe s'appelle *courbe du second degré* ou *conique*. En fait il peut n'exister aucun point $(x, y)$

de coordonnées réelles vérifiant l'équation (1). Dans ce cas on dit que l'équation (1) définit une *conique imaginaire.* Nous glisserons sur l'étude des coniques imaginaires. L'équation

$$x^2 + y^2 = -1$$

définit une conique imaginaire.

Citons les six cas particuliers les plus importants de l'équation (1).

1) Equation de *l'ellipse*:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a \geqslant b > 0),$$

$2a$ et $2b$ étant les longueurs des grand et petit axes. Si $a = b$, on obtient l'équation du *cercle*:

$$x^2 + y^2 = a^2,$$

de centre $O$ et de rayon $a$.

2) Equation de l'*hyperbole*:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (a \geqslant b > 0),$$

$2a$ et $2b$ étant les longueurs des grand et petit axes.

3) Equation de la *parabole*:

$$y^2 = 2px \quad (p > 0).$$

4) Equation d'un *couple de droites concourantes*:

$$a^2x^2 - b^2y^2 = 0 \quad (0 < a, \ b).$$

5) Equation d'un *couple de droites parallèles* ou *confondues*:

$$x^2 - a^2 = 0 \quad (a \geqslant 0).$$

6) Equation d'un *point*:

$$x^2 + y^2 = 0.$$

Etudions brièvement les courbes citées.

L'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (a \geqslant b > 0). \tag{2}$$

Pour $a = b$, l'ellipse (2) se transforme en un cercle de rayon $a$ et de centre $O$, c'est-à-dire le lieu géométrique des points situés à une distance constante $a$ de l'origine.

Supposons que $a > b$. Posons $c = \sqrt{a^2 - b^2}$. Les points $F_1$ et $F_2$ de l'axe $Ox$, d'abscisses respectives $x = -c$ et $x = c$ s'appellent *foyers de l'ellipse.* L'ellipse (2) se définit comme le lieu géométrique des points dont la somme des distances aux foyers $F_1$ et $F_2$ est une constante égale à $2a$.

En effet (fig. 145)

$$MF_1 = \sqrt{(x+c)^2+y^2}, \quad MF_2 = \sqrt{(x-c)^2+y^2},$$
$$2a = \sqrt{(x+c)^2+y^2} + \sqrt{(x-c)^2+y^2},$$

d'où

$$2a - \sqrt{(x+c)^2+y^2} = \sqrt{(x-c)^2+y^2}$$

et

$$4a^2 + (x+c)^2 + y^2 - 4a\sqrt{(x+c)^2+y^2} = (x-c)^2 + y^2,$$
$$4a^2 + 4cx = 4a\sqrt{(x+c)^2+y^2},$$
$$a^4 + 2a^2cx + c^2x^2 = a^2\,[x^2 + 2cx + c^2 + y^2],$$
$$-b^2x^2 = -a^2b^2 + a^2y^2,$$
$$a^2b^2 = b^2x^2 + a^2y^2,$$

d'où l'on déduit l'équation (2). Si l'on remonte ces calculs, on trouve que si un point $(x, y)$ vérifie l'équation (2), la somme de ses distances à $F_1$ et $F_2$ est égale à $2a$.

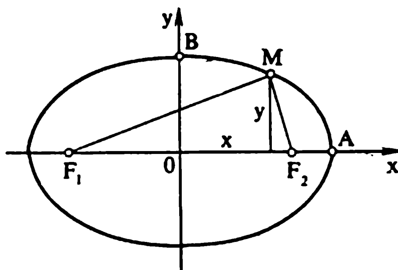



Fig. 145

Si dans l'équation (2) on remplace $x$ par $-x$, on constate que cette équation ne change pas. Ceci exprime que l'ellipse (2) est une courbe symétrique par rapport à l'axe $Oy$. On vérifie immédiatement qu'elle est symétrique aussi par rapport à l'axe $Ox$. On peut donc étudier son équation dans le premier quadrant seulement, c'est-à-dire pour $x, y \geqslant 0$. La portion d'ellipse située dans le premier quadrant est définie par l'équation:

$$y = \frac{b}{a}\sqrt{a^2 - x^2}, \qquad 0 \leqslant x \leqslant a.$$

On remarque que l'ellipse passe par les points $(0, b)$ et $(a, 0)$, et que, lorsque $x$ croît continûment sur l'intervalle $[0, a]$, l'ordonnée $y$ décroît continûment.

L'ellipse est une courbe bornée. Elle est comprise dans un cercle de rayon $a$ et de centre $O$ (les coordonnées de tout point de l'ellipse vérifient l'inégalité $x^2 + y^2 \leqslant a^2 \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right) = a^2$).

Sur la figure 145, on voit que l'ellipse est une courbe continue fermée. Dans le premier quadrant c'est une courbe convexe vers le haut. Elle admet une tangente [1] en chacun de ses points. Toutes ces propriétés et beaucoup d'autres peuvent être étudiées par les méthodes d'analyse mathématique. L'analyse mathématique nous fournit du reste des outils pour la définition rigoureuse des notions de continuité, de convexité et autres, évoquées plus haut.

L'équation de l'ellipse peut se mettre sous la forme paramétrique:

$$\begin{cases} x = a\cos\theta, \\ y = b\sin\theta \end{cases} \quad (-\infty < \theta < \infty). \tag{3}$$

En effet,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \cos^2\theta + \sin^2\theta = 1,$$

c'est-à-dire que tout point $(x, y)$ défini par les égalités (3) appartient à l'ellipse (2) quel que soit $\theta$. Si $\theta$ parcourt de façon continue l'intervalle $[0, 2\pi[$, alors le point $(x, y)$ décrit l'ellipse tout entière. Si $\theta$ continue de croître, le mouvement se répète périodiquement.

Voyons quelle est la signification du paramètre $\theta$ et indiquons incidemment une méthode de construction de l'ellipse (fig. 146). Traçons deux cercles concentriques de rayons $b$ et $a$ ($b < a$) et de centre $O$. Menons ensuite un rayon vecteur faisant un angle $\theta$ avec l'axe $Ox$ et désignons par $T$ et $N$ ses points d'intersection respectifs avec les cercles de rayons $b$ et $a$. La droite parallèle à l'axe $Oy$, qui passe par $N$, coupe la droite parallèle à l'axe $Ox$, qui passe par $T$, en un point $M$ qui appartient à l'ellipse. En effet, soient $x$, $y$ les coordonnées du point $M$. Alors (fig. 146)

$$x = ON\cdot\cos\theta = a\cos\theta,$$
$$y = TR = OT\cdot\sin\theta = b\sin\theta,$$

c'est-à-dire que le point $M$ appartient bien à l'ellipse (3), et le paramètre $\theta$ est l'angle de l'axe $Ox$ et du rayon $ON$. Signalons que $\theta$ n'est pas l'angle polaire $\varphi$ du rayon vecteur $OM$ avec l'axe $Ox$ $\left(\operatorname{tg}\varphi = \frac{b}{a}\operatorname{tg}\theta\right)$. Par exemple, si $\varphi = \pi/4$, $a = \sqrt{3}$, $b = 1$, alors $\theta = \pi/3$; si $\varphi = 0$, alors $\theta = 0$; si $\varphi = \pi/2$, alors $\theta = \pi/2$.

L'hyperbole

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \qquad (0 < a, b). \tag{4}$$

Posons $c = \sqrt{a^2 + b^2}$. Les points $F_1$ et $F_2$ de l'axe $Ox$, d'abscisses respectives $x = -c$ et $x = c$ (fig. 147), s'appellent *foyers de l'hyperbole* (4).

[1] Voir chap. 4, § 2.

L'hyperbole (4) se définit aussi comme le lieu géométrique des points $A = (x, y)$ dont la différence des distances aux foyers $F_1$ et $F_2$ est une grandeur constante égale à $2a$.

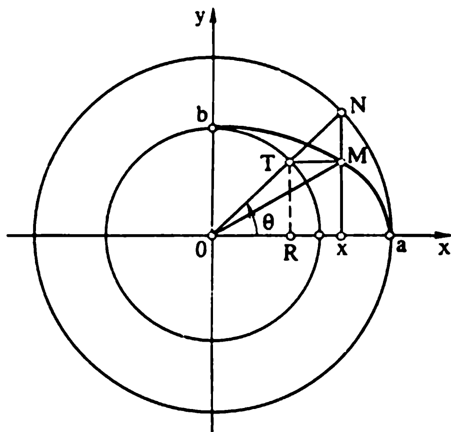



Fig. 146

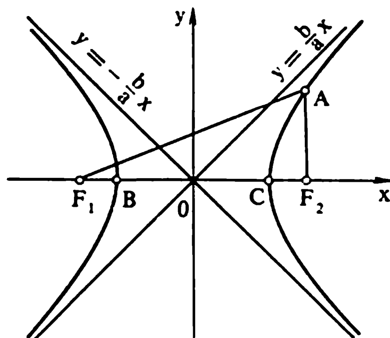



Fig. 147

On a (voir fig. 147)

$$AF_1 - AF_2 = \sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = 2a,$$

$$\sqrt{(x+c)^2 + y^2} = \sqrt{(x-c)^2 + y^2} + 2a,$$

$$(x+c)^2 + y^2 = (x-c)^2 + y^2 + 4a\sqrt{(x-c)^2 + y^2} + 4a^2,$$

$$4cx - 4a^2 = 4a\sqrt{(x-c)^2 + y^2},$$

$$c^2x^2 - 2a^2cx + a^4 = a^2(x^2 - 2cx + c^2) + a^2y^2,$$

$$(a^2 + b^2)x^2 = a^2x^2 + a^2b^2 + a^2y^2,$$

$$b^2x^2 - a^2y^2 = a^2b^2,$$

d'où l'équation (4).

Nous avons obtenu la branche droite de l'hyperbole (fig. 147). Pour obtenir la branche gauche il faut se servir de l'égalité

$$AF_2 - AF_1 = 2a.$$

En remontant ces calculs on montre que tout point $(x, y)$ vérifiant l'équation (4) appartient à ce lieu géométrique.

La forme de l'équation (4) nous dit que l'hyperbole est une courbe symétrique par rapport aux axes $Ox$ et $Oy$. L'équation de la portion d'hyperbole comprise dans le premier quadrant est

$$y = \frac{b}{a}\sqrt{x^2 - a^2} \qquad (a \leqslant x < \infty). \tag{5}$$

On voit que l'hyperbole passe par le point $(a, 0)$ et que $y$ croît et tend vers l'infini lorsque $x$ croît sur l'intervalle $[a, \infty[$. Les points $B = (-a, 0)$ et $C = (a, 0)$ d'intersection de l'hyperbole et de l'axe $Ox$ s'appellent *sommets de l'hyperbole.*

Les deux droites:

$$y = \pm \frac{b}{a} x,$$

représentées sur la figure 147, sont les *asymptotes* de l'hyperbole.

Soit donnée une courbe $y = f(x)$ sur l'intervalle $[a, \infty[$ (ou $]-\infty, a]$). On dit qu'une droite $y = mx + n$ est une *asymptote* de cette courbe pour $x \to +\infty$ (resp. $x \to -\infty$) si

$$\lim_{x \to +\infty} [f(x) - mx - n] = 0$$

(resp. $\lim\limits_{x \to -\infty} [f(x) - mx - n] = 0$).

Considérons la portion d'hyperbole définie par l'équation (5) et comparons-la avec la droite $y = \frac{b}{a} x$. On a

$$\lim_{x \to +\infty} \left[\frac{bx}{a} - \frac{b}{a}\sqrt{x^2 - a^2}\right] = \lim_{x \to +\infty} \frac{b}{a} \frac{a^2}{x + \sqrt{x^2 - a^2}} = 0.$$

Ceci montre que la droite $y = \frac{b}{a} x$ est une asymptote de la portion d'hyperbole considérée pour $x \to +\infty$. On dit alors que la droite $y = \frac{b}{a} x$ est l'asymptote de (toute!) l'hyperbole pour $x \to +\infty$. L'hyperbole étant symétrique par rapport aux axes de coordonnées, de même d'ailleurs que le couple de droites $y = \pm \frac{b}{a} x$, on peut dire que ces droites sont les asymptotes de l'hyperbole, aussi bien pour $x \to +\infty$ que pour $x \to -\infty$.

La branche droite de l'hyperbole (4) peut se mettre sous la forme paramétrique:

$$\begin{cases} x = a \operatorname{ch} u = \frac{a}{2}(e^u + e^{-u}), \\ y = b \operatorname{sh} u = \frac{b}{2}(e^u - e^{-u}) \end{cases} \quad (u \in ]-\infty, \infty[). \tag{6}$$

En effet, comme

$$\operatorname{ch}^2 u - \operatorname{sh}^2 u = 1, \tag{7}$$

on déduit des équations (6)

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \operatorname{ch}^2 u - \operatorname{sh}^2 u = 1.$$

La moitié supérieure de la branche droite de l'hyperbole correspond à une variation de $u$ sur l'intervalle $[0, \infty[$, la moitié inférieure, à une variation sur l'intervalle $]-\infty, 0]$.

Voyons comment le paramètre $u$ est relié au paramètre $\theta$ de l'équation de l'ellipse et indiquons incidemment une méthode de construction de l'hyperbole à la règle et au compas. Notre méthode de construction de l'hyperbole étant basée sur celle de l'ellipse, nous allons exposer ces méthodes simultanément (fig. 148). On se bornera à construire la portion de l'ellipse (2) et de l'hyperbole (6) délimitée par le premier quadrant. Traçons deux cercles concentriques de rayons $a$ et $b$ centrés en l'origine des coordonnées. Appelons $T_1$ et $N_1$ les

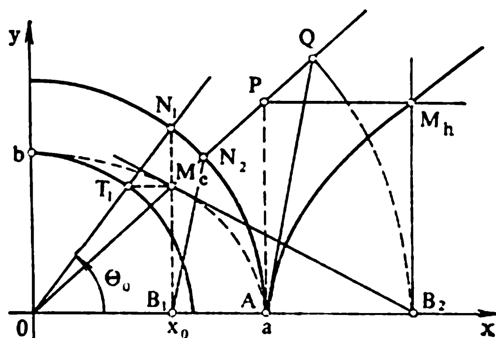



Fig. 148

points d'intersection de la droite issue de $O$ sous un angle $\theta_0$ avec l'axe $Ox$ ($OT_1 = b$, $ON_1 = a$). Les parallèles menées à partir de $T_1$ et de $N_1$ respectivement à l'axe $Ox$ et à l'axe $Oy$ se coupent en un point $M_e = (x_0, y_0)$ appartenant à l'ellipse (2). Soient $N_2$ le point d'intersection de $OM_e$ et du cercle de rayon $a$; $P$, le point d'intersection de $OM_e$ avec la droite parallèle à l'axe $Oy$, qui passe par le point $A = (a, 0)$ de l'ellipse. L'équation de $OP$ est

$$y = \frac{y_0}{x_0} x.$$

Il s'ensuit que l'ordonnée du point $P$ est $Y_0 = \frac{ay_0}{x_0}$. Joignons le point $B_1 = (x_0, 0)$ et le point $N_2$, et à partir de $A$ menons une droite parallèle à $B_1N_2$ jusqu'à son intersection $Q$ avec $OP$. La similitude des triangles $OAQ$ et $OB_1N_2$ nous donne $OQ = a^2/x_0$. Traçons un arc de cercle de rayon $OQ$ jusqu'à son intersection $B_2 = (a^2/x_0, 0)$ avec l'axe $Ox$.

A partir des points $B_2$ et $P$ menons des droites parallèles respectivement à $Oy$ et $Ox$. Le point d'intersection $M_h = (X_0, Y_0)$, où $X_0 = a^2/x_0$, appartient à l'hyperbole (4).

En effet, comme le point $(x_0, y_0)$ est situé sur l'ellipse (2), on a

$$\frac{X_0^2}{a^2} - \frac{Y_0^2}{b^2} = \frac{a^4}{a^2x_0^2} - \frac{a^2y_0^2}{x_0^2b^2} = \frac{a^2}{x_0^2}\left(1 - \frac{y_0^2}{b^2}\right) = \frac{a^2}{x_0^2} \cdot \frac{x_0^2}{a^2} = 1,$$

c'est-à-dire que le point $M_h$ appartient à l'hyperbole (4).

Signalons que le point $B_2 = (a^2/x_0, 0)$ est le point de concours de la tangente à l'ellipse en $M_e$ avec l'axe $Ox$ (voir note de la page 476).

Donc, à tout point $(x, y)$ de l'ellipse (2) est associé un point bien défini $(X, Y)$ de l'hyperbole (4) et réciproquement.

Si maintenant l'ellipse (2) est donnée sous forme paramétrique, alors

$$x = a\cos\theta, \quad y = b\sin\theta.$$

Donc

$$X = \frac{a^2}{x} = \frac{a}{\cos\theta}, \quad Y = \frac{ay}{x} = b\,\mathrm{tg}\,\theta.$$

D'où, compte tenu de (6), on déduit

$$\mathrm{ch}\,u = \frac{1}{\cos\theta}, \quad \mathrm{sh}\,u = \mathrm{tg}\,\theta.$$

On a les formules suivantes :

$$\mathrm{tg}\,\frac{\theta}{2} = \sqrt{\frac{1-\cos\theta}{1+\cos\theta}} = \sqrt{\frac{\mathrm{ch}\,u-1}{\mathrm{ch}\,u+1}} = \mathrm{th}\,\frac{u}{2};$$

$$e^u = \mathrm{ch}\,u + \mathrm{sh}\,u = \frac{1}{\cos\theta} + \mathrm{tg}\,\theta = \frac{1+\sin\theta}{\cos\theta} =$$

$$= \frac{\left(\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\right)^2}{\cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}} = \frac{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}}{\cos\frac{\theta}{2} - \sin\frac{\theta}{2}} = \frac{\sin\left(\frac{\pi}{4} + \frac{\theta}{2}\right)}{\cos\left(\frac{\pi}{4} + \frac{\theta}{2}\right)} = \mathrm{tg}\left(\frac{\pi}{4} + \frac{\theta}{2}\right),$$

c'est-à-dire que

$$u = \ln \mathrm{tg}\left(\frac{\pi}{4} + \frac{\theta}{2}\right).$$

L a   p a r a b o l e

$$y^2 = 2px \qquad (p > 0). \tag{8}$$

Le point $F$ de l'axe $Ox$ d'abscisse $x = p/2$ s'appelle *foyer de la parabole* (8). La droite $x = -p/2$ s'appelle *directrice de la parabole* (fig. 149).

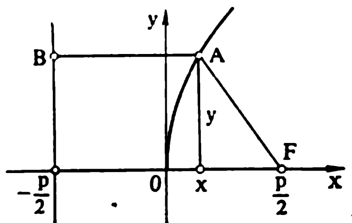



Fig. 149

La parabole (8) peut encore être définie comme le lieu géométrique des points $A = (x, y)$ équibistants du foyer et de la directrice.

En effet (fig. 149)

$$AF^2 = \left(x - \frac{p}{2}\right)^2 + y^2,$$

$$AB^2 = \left(x + \frac{p}{2}\right)^2,$$

donc

$$\left(x - \frac{p}{2}\right)^2 + y^2 = \left(x + \frac{p}{2}\right)^2$$

$$-px + y^2 = px,$$

i.e.

$$y^2 = 2px.$$

Réciproquement, les points dont les coordonnées vérifient cette équation appartiennent au lieu géométrique indiqué. On voit de

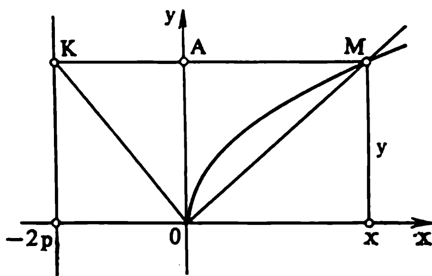



Fig. 150

l'équation (8) que la parabole est symétrique par rapport à l'axe $Ox$. L'équation de sa moitié supérieure est

$$y = \sqrt{2px} \quad (0 \leqslant x < \infty). \tag{9}$$

On remarque que $y$ tend vers l'infini avec $x$.

Voyons une méthode simple de construction de la parabole (9) à la règle et l'équerre ou à la règle et au compas. Traçons la droite $x = -2p$ (fig. 150). Prenons sur cette droite un point quelconque $K = (-2p, y)$, $y > 0$. Traçons $OK$ et une droite perpendiculaire à $OK$ en $O$. Menons ensuite par $K$ une droite parallèle à $Ox$. Les deux dernières droites se coupent en un point $M = (x, y)$ situé sur la parabole (9), puisque $OA = y$ est la moyenne géométrique des nombres $2p$ et $x$ ($y = \sqrt{2px}$).

La parabole (8) n'admet pas d'asymptote [1]).

Le couple de droites concourantes

$$a^2x^2 - b^2y^2 = (ax - by)(ax + by) = 0 \quad (0 < a,\ b). \tag{10}$$

[1]) Voir chap. 4, § 20.

Si un point $(x, y)$ vérifie l'équation (10), il vérifiera l'une des équations

$$\begin{cases} ax - by = 0, \\ ax + by = 0 \end{cases} \tag{10'}$$

ou les deux. Réciproquement, si un point $(x, y)$ vérifie l'une des équations (10'), il vérifiera l'équation (10). On dit que l'équation (10) est l'équation d'un couple de droites.

On démontre plus bas qu'il existe un système de coordonnées rectangulaires dans lequel la courbe (1), sous réserve qu'elle ne soit pas imaginaire, a pour équation l'une des équations 1) à 6) énumérées plus haut.

De façon plus détaillée:

si $AC - B^2 > 0$, la courbe (1) est une ellipse, un point (cas 1) et 6)) ou une courbe imaginaire;

si $AC - B^2 < 0$, la courbe (1) est une hyperbole ou un couple de droites concourantes (cas 2) et 4));

si $AC - B^2 = 0$, la courbe (1) est une parabole, un couple de droites parallèles ou confondues ou une courbe imaginaire (cas 3) et 5)).

Par abus de langage on parlera de « courbe » même dans les cas 4), 5), 6), où il est question d'un couple de droites ou d'un ensemble constitué d'un seul point.

Ainsi, soit donnée l'équation

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \tag{1}$$

où les coefficients $A$, $B$ et $C$ ne sont pas tous nuls.

Sans nuire à la généralité on peut admettre que $A \geqslant 0$, $A \geqslant C$, $B \geqslant 0$. On peut toujours se ramener à ce cas par les transformations orthogonales:

$$\begin{cases} x = \eta, \\ y = \xi, \end{cases} \qquad \begin{cases} x = -\xi, \\ y = \eta \end{cases}$$

et en multipliant les deux membres de (1) par $-1$.

Si $B = 0$, $A \geqslant C > 0$, on peut écrire (1) sous la forme

$$A\left(x + \frac{D}{A}\right)^2 + C\left(y + \frac{E}{C}\right)^2 + F - \frac{E^2}{C} - \frac{D^2}{A} = 0. \tag{11}$$

La translation

$$\xi = x + \frac{D}{A}, \qquad \eta = y + \frac{E}{C}$$

transforme l'équation (11) en

$$A\xi^2 + C\eta^2 = \frac{E^2}{C} + \frac{D^2}{A} - F. \tag{11'}$$

Si $\frac{E^2}{C}+\frac{D^2}{A}-F>0$, l'équation (11') décrit une ellipse d'axes $2a$, $2b$, où

$$a^2=\left(\frac{E^2}{C}+\frac{D^2}{A}-F\right)\Big/A, \qquad b^2=\left(\frac{E^2}{C}+\frac{D^2}{A}-F\right)\Big/C.$$

Signalons qu'ici $AC-B^2=AC>0$.

Si le second membre de (11') est nul, on obtient un point (cas 6)).

Si le second membre de (11') est strictement négatif, on obtient une courbe imaginaire.

Pour $C=0$, $A>0$, l'équation (1) devient

$$A\left(x+\frac{D}{A}\right)^2+2Ey+F-\frac{D^2}{A}=0. \tag{12}$$

Si $E\neq 0$, la translation

$$\xi=x+\frac{D}{A}, \qquad \eta=y+\frac{F}{2E}-\frac{D^2}{2AE}$$

transforme l'équation (12) en l'équation

$$A\xi^2+2E\eta=0 \tag{12'}$$

qui (quitte à substituer $\eta$ à $-\eta$) décrit une parabole (cas 3)).

Si $E=0$, on obtient un couple de droites parallèles ou une courbe imaginaire selon que $\frac{D^2}{A}-F$ est strictement positif ou négatif. Signalons qu'ici $AC-B^2=0$.

Si $C<0$, $A>0$, alors l'équation (1) devient

$$A\left(x+\frac{D}{A}\right)^2-|C|\left(y-\frac{E}{|C|}\right)^2+F+\frac{E^2}{|C|}-\frac{D^2}{A}=0, \tag{13}$$

équation qui se discute de même façon que l'équation (11). L'équation (13) décrit une hyperbole ou un couple de droites concourantes (cas 2) et 4)). Signalons qu'ici $AC-B^2=AC<0$.

Le cas $A=0$, $C<0$ se ramène à une équation de la forme (12).

Donc, pour $B=0$, l'équation (1) nous ramène toujours à l'un des cas particuliers 1) à 6).

Supposons que $B>0$, $A\geqslant C$. On sait (voir § 23) qu'il existe une transformation orthogonale

$$\begin{cases} x=x_1\xi-y_1\eta, \\ y=y_1\xi+x_1\eta, \end{cases} \tag{14}$$

où

$$x_1=\sqrt{\frac{1}{2}+\frac{A-C}{2\sqrt{4B^2+(A-C)^2}}},$$
$$y_1=\sqrt{\frac{1}{2}-\frac{A-C}{2\sqrt{4B^2+(A-C)^2}}},$$

qui réduit la forme quadratique

$$Ax^2 + 2Bxy + Cy^2$$

à sa forme canonique.

Transformons l'équation (1) à l'aide de (14):

$$\lambda_1\xi^2 + \lambda_2\eta^2 + D(x_1\xi - y_1\eta) + E(y_1\xi + x_1\eta) + F = 0, \qquad (15)$$

où

$$\lambda_1 = \frac{1}{2}\left[A + C + \sqrt{4B^2 + (A-C)^2}\right],$$

$$\lambda_2 = \frac{1}{2}\left[A + C - \sqrt{4B^2 + (A-C)^2}\right]$$

($\lambda_1 > \lambda_2$, $\lambda_1\lambda_2 = AC - B^2$). Mettons l'équation (15) sous la forme

$$\lambda_1\xi^2 + \lambda_2\eta^2 + (x_1D + y_1E)\,\xi + (x_1E - y_1D)\,\eta + F = 0. \qquad (15')$$

L'équation (15′) est un cas particulier de l'équation (1) pour $B = 0$, cas que nous avons déjà étudié.

On peut donc dire que si:

1) $AC - B^2 = \lambda_1\lambda_2 > 0$, alors l'équation (1) décrit une ellipse, un point ou une courbe imaginaire. Dans ce cas, on dira que l'équation (1) est de *type elliptique*;

2) $AC - B^2 = \lambda_1\lambda_2 < 0$, alors l'équation (1) décrit une hyperbole ou un couple de droites concourantes. On dira alors que l'équation (1) est de *type hyperbolique*;

3) $AC - B^2 = \lambda_1\lambda_2 = 0$, alors l'équation (1) décrit une parabole, un couple de droites parallèles ou une courbe imaginaire. On dit que l'équation (1) est de *type parabolique*.

Exemple.. Nature de la courbe

$$2x^2 + \sqrt{3}\,xy + y^2 + 2x + 2\sqrt{3}\,y + F = 0,$$

où $F$ est un nombre réel.

On a $A = 2 > C = 1$, $B = \frac{\sqrt{3}}{2} > 0$, $AC - B^2 = \frac{5}{4} > 0$, donc l'équation est de type elliptique. On établit sans peine (voir exemple du § 23) que

$$x_1 = \frac{\sqrt{3}}{2}, \quad y_1 = \frac{1}{2}, \quad \lambda_1 = \frac{5}{2}, \quad \lambda_2 = \frac{1}{2}.$$

Donc, la transformation orthogonale

$$x = \frac{1}{2}(\sqrt{3}\,\xi - \eta),$$

$$y = \frac{1}{2}(\xi + \sqrt{3}\,\eta)$$

réduit l'équation donnée à la forme

$$\frac{5}{2}\xi^2 + \frac{1}{2}\eta^2 + (\sqrt{3}\,\xi - \eta) + \sqrt{3}(\xi + \sqrt{3}\,\eta) + F = 0$$

ou encore

$$\frac{5}{2}\left(\xi+\frac{2}{5}\sqrt{3}\right)^2+\frac{1}{2}(\eta+2)^2=\frac{16}{5}-F.$$

La translation

$$u=\xi+\frac{2}{5}\sqrt{3},$$
$$v=\eta+2$$

nous conduit à l'équation

$$\frac{5}{2}u^2+\frac{1}{2}v^2=\frac{16}{5}-F. \tag{16}$$

Si $\frac{16}{5}-F>0$, alors (16) est l'équation d'une ellipse d'axes $2a$, $2b$, où

$$a^2=2(16-5F)/25, \quad b^2=2(16-5F)/5.$$

Si $\frac{16}{5}-F=0$, alors (16) est l'équation d'un point. Si $\frac{16}{5}-F<0$, l'équation (16) décrit une courbe imaginaire.

## § 25. Quadriques dans l'espace

L'équation

$$\sum_{k=1}^{3}\sum_{l=1}^{3}a_{kl}x_kx_l+2\sum_{l=1}^{3}A_lx_l+B=0, \tag{1}$$

où $a_{kl}=a_{lk}$, $A_l$, $B$ sont des constantes données, $x=(x_1, x_2, x_3)$ un point variable dans $R_3$, définit un ensemble de points dans $R_3$ appelé *surface du second degré* ou *quadrique*. Si l'équation (1) n'est vérifiée par aucun point réel $x=(x_1, x_2, x_3)$, on dit qu'elle décrit une *surface imaginaire*. Nous passerons sur l'étude de tels cas. L'équation (1) définit parfois un couple de plans distincts ou confondus ou un point unique. Ces ensembles seront appelés aussi surfaces.

Citons les principaux cas particuliers de l'équation (1):

1) *L'ellipsoïde*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \quad (a, b, c>0).$$

2) *L'hyperboloïde à une nappe*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a, b, c>0).$$

3) *L'hyperboloïde à deux nappes*

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a, b, c>0).$$

4) *Le paraboloïde elliptique*

$$\frac{x^2}{p}+\frac{y^2}{q}=2z \quad (p,\ q>0).$$

5) *Le paraboloïde hyperbolique*

$$\frac{x^2}{p}-\frac{y^2}{q}=2z \quad (p,\ q>0).$$

6) *Le cône du second degré*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \quad (a,\ b,\ c>0).$$

7) *Le point*

$$x^2+y^2+z^2=0.$$

8) *Les cylindres du second degré :*

— *le cylindre elliptique*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad (a,\ b>0),$$

— *le cylindre hyperbolique*

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (a,\ b>0),$$

— *le cylindre parabolique*

$$y^2=2px \quad (p>0),$$

— *le couple de plans concourants*

$$a^2x^2-b^2y^2=0 \quad (a,\ b>0),$$

— *le couple de plans parallèles ou confondus*

$$x^2-a^2=0 \quad (a>0),$$

$$z^2=0,$$

— *la droite*

$$x^2+y^2=0.$$

On a admis que $x=x_1$, $y=x_2$, $z=x_3$.

On démontre que chaque cas particulier de l'équation (1), sous réserve qu'elle ne soit pas l'équation d'une surface imaginaire, définit dans un certain système de coordonnées rectangulaires l'une des 8 surfaces énumérées. Ceci résulte de la théorie générale du § 22. L'équation (1) se transforme comme dans le paragraphe 24. La détermination des valeurs propres $\lambda_1$, $\lambda_2$, $\lambda_3$ se ramène à la résolution d'une équation du troisième degré.

Indiquons encore une méthode de détermination des valeurs et vecteurs propres, méthode dont le principe a déjà été développé au § 23 dans le plan. Les valeurs propres $\lambda_1$, $\lambda_2$, $\lambda_3$ de l'opérateur hermitien $A$ et les vecteurs unités propres correspondants $x^1$, $x^2$, $x^3$

($|x^j| = 1$, $Ax^j = \lambda_j x^j$, $j = 1, 2, 3$) se déterminent de la manière suivante (voir justification plus bas). Considérons le déterminant ($a_{kl} = a_{lk}$)

$$D(\lambda) = |A - \lambda E| = \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix},$$

où $E$ est la matrice unité. Trouvons les racines $\lambda_1$, $\lambda_2$, $\lambda_3$ de l'équation

$$|A - \lambda E| = 0, \qquad (2)$$

qui s'appelle *équation caractéristique* de l'opérateur $A$ ($D(\lambda_j) = 0$, $j = 1, 2, 3$). Ces racines ne sont autres que les valeurs propres de l'opérateur $A$. Elles sont réelles et peuvent être distinctes ou confondues (multiples). Donc

$$D(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda).$$

On cherche ensuite la solution triviale $x = (x_1, x_2, x_3)$ du système d'équations homogène, correspondant à $\lambda_1$:

$$\begin{cases} (a_{11}-\lambda_1)x_1 + a_{12}x_2 + a_{13}x_3 = 0, \\ a_{12}x_1 + (a_{22}-\lambda_1)x_2 + a_{23}x_3 = 0, \\ a_{31}x_1 + a_{32}x_2 + (a_{33}-\lambda_1)x_3 = 0 \end{cases} \qquad (3)$$

ou de l'équation correspondante pour l'opérateur $A - \lambda_1 E$:

$$(A - \lambda_1 E)\, x = 0, \qquad (3')$$

où $E$ est la matrice unité et $x = (x_1, x_2, x_3)$ et $0 = (0, 0, 0)$.

Si $\lambda_1$ est une racine simple (c'est-à-dire distincte de $\lambda_2$ et de $\lambda_3$), le rang de la matrice du système (3) doit nécessairement être égal à 2 (rang $(A - \lambda_1 E) = 2$) et on obtient un vecteur unique, au signe près, $x^1 = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$ vérifiant le système (3), c'est-à-dire que

$$(A - \lambda_1 E)\, x^1 = 0$$

ou

$$Ax^1 = \lambda_1 x^1.$$

Si $\lambda_1$ est une racine double ($\lambda_1 = \lambda_2 \neq \lambda_3$), alors le rang de la matrice du système (3) est nécessairement égal à 1 (rang $(A - \lambda_1 E) = 1$) et l'on aura pour solutions deux vecteurs unités orthogonaux $x^1$ et $x^2$ ($|x^1| = |x^2| = 1$, $(x^1, x^2) = 0$) qui sont les vecteurs propres associés à la valeur propre $\lambda_1$:

$$Ax^j = \lambda_j x^j \quad (j = 1, 2, \ \lambda_1 = \lambda_2).$$

Enfin, si $\lambda_1$ est une racine de multiplicité trois ($\lambda_1 = \lambda_2 = \lambda_3$), alors le rang de la matrice du système (3) est nécessairement nul (rang $(A - \lambda_1 E) = 0$). Ce système admet donc pour solutions trois

vecteurs unités orthogonaux $x^1$, $x^2$, $x^3$:

$$Ax^j = \lambda_j x^j \quad (j=1,\ 2,\ 3,\ \lambda_1 = \lambda_2 = \lambda_3).$$

Pour vecteurs propres $x^1$. $x^2$, $x^3$ associés aux valeurs propres $\lambda_1 = \lambda_2 = \lambda_3$ on peut prendre trois vecteurs unités orthogonaux quelconques de $R_3$.

Justifions ce qui vient d'être dit. On sait que dans $R_3$ il existe un système orthonormal de vecteurs $x^1$, $x^2$, $x^3$ et des nombres réels $\lambda_1$, $\lambda_2$, $\lambda_3$, tels que

$$Ax^j = \lambda_j x^j \quad (j = 1,\ 2,\ 3).$$

De plus, on peut exhiber une matrice orthogonale $\Lambda$ telle que (voir § 22)

$$\Lambda A \Lambda^{-1} = \begin{Vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{Vmatrix}.$$

D'où l'on déduit l'identité pour la variable $\lambda$:

$$\Lambda (A - \lambda E) \Lambda^{-1} = \begin{Vmatrix} \lambda_1 - \lambda & 0 & 0 \\ 0 & \lambda_2 - \lambda & 0 \\ 0 & 0 & \lambda_3 - \lambda \end{Vmatrix} \tag{4}$$

$$(\Lambda (A - \lambda E) \Lambda^{-1} = \Lambda A \Lambda^{-1} - \Lambda \lambda E \Lambda^{-1} = \\ = \Lambda A \Lambda^{-1} - \lambda \Lambda \Lambda^{-1} = \Lambda A \Lambda^{-1} - \lambda E)$$

On désigne toujours le déterminant de la matrice $A - \lambda E$ par $D(\lambda)$. De (4) on voit que $D(\lambda)$ est égal à celui de la matrice du second membre de (4), puisque

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) = |\Lambda (A - \lambda E) \Lambda^{-1}| = \\ = |\Lambda|\,|A - \lambda E|\,|\Lambda^{-1}| = |A - \lambda E| = D(\lambda).$$

Donc

$$D(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda).$$

On obtient que les racines du polynôme $D(\lambda)$ sont confondues avec les valeurs propres $\lambda_1$, $\lambda_2$, $\lambda_3$ de l'opérateur $A$. Ces racines sont par conséquent *réelles*.

Supposons que $\lambda_1$ est une racine simple, donc $\lambda_1 \neq \lambda_2$ et $\lambda_1 \neq \lambda_3$. Lorsque $\lambda = \lambda_1$, le rang de la matrice du second membre de (4) est égale à 2 (rang $\Lambda (A - \lambda_1 E) \Lambda^{-1} = 2$), donc rang $(A - \lambda_1 E) = 2$. En effet, les solutions du système homogène (3) et du système

$$\begin{cases} 0 \cdot z_1 + 0 \cdot z_2 + 0 \cdot z_3 = 0, \\ 0 \cdot z_1 + (\lambda_2 - \lambda_1) z_2 + 0 \cdot z_3 = 0, \\ 0 \cdot z_1 + 0 \cdot z_2 + (\lambda_3 - \lambda_1) z_3 = 0 \end{cases} \tag{5}$$

associé à la matrice (4) se transforment l'une en l'autre à l'aide d'une matrice orthogonale (les systèmes (3) et (5) sont équivalents). Le système (5) n'admet que la solution $(\pm 1, 0, 0)$ au signe près. Or, ceci n'est possible que si rang $(A - \lambda_1 E) = 2$.

On peut encore justifier ce fait de la manière suivante. Si l'on admet que tous les déterminants d'ordre deux de la matrice $A - \lambda_1 E$ sont nuls, alors tous les déterminants d'ordre deux de la matrice $\Lambda (A - \lambda_1 E) \Lambda^{-1}$ seront aussi nuls, puisqu'ils sont des combinaisons linéaires des déterminants d'ordre deux de la matrice $A - \lambda_1 E$. Mais ceci est impossible, car le déterminant de la matrice $\Lambda (A - - \lambda_1 E) \Lambda^{-1}$

$$\begin{vmatrix} \lambda_2 - \lambda_1 & 0 \\ 0 & \lambda_3 - \lambda_1 \end{vmatrix} \neq 0.$$

Si maintenant $\lambda_1 = \lambda_2 \neq \lambda_3$, par des raisonnements analogues on établit que rang $(A - \lambda_1 E) = 1$ et le système (3) admet pour solutions exactement deux vecteurs unités orthogonaux $x^1$, $x^2$ associés à $\lambda_1 = \lambda_2$.

Enfin, pour $\lambda_1 = \lambda_2 = \lambda_3$ on a rang $(A - \lambda_1 E) = 0$, c'est-à-dire que tous les éléments de la matrice $A - \lambda_1 E$ sont nuls. Dans ce cas, tout vecteur $x = (x_1, x_2, x_3)$ est solution du système (3). Pour vecteurs propres associés à la valeur propre $\lambda_1 = \lambda_2 = \lambda_3$ on peut alors prendre trois vecteurs quelconques $x^1$, $x^2$, $x^3$ formant un système orthonormal.

On remarquera que dans ce cas la forme quadratique est déjà réduite à une somme de carrés $(a_{12} = a_{13} = a_{23} = 0,\ a_{11} = a_{22} = = a_{33} = \lambda_1)$.

Exemple. Réduire la forme quadratique

$$x_1^2 + x_2^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3.$$

Ici $a_{11} = a_{22} = a_{12} = a_{13} = a_{23} = 1,\ a_{33} = 0$. Formons l'équation caractéristique:

$$\begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = 0$$

ou

$$-\lambda (1 - \lambda)^2 + 2 - 2 (1 - \lambda) + \lambda = 0, \quad -\lambda (1 - \lambda)^2 + 3\lambda = 0.$$

Il est immédiat que $\lambda = 0$ est une racine de cette équation. Trouvons les deux autres:

$$3-(1-\lambda)^2 = 0, \quad (1-\lambda)^2 = 3, \quad 1-\lambda = \pm\sqrt{3}, \quad \lambda = 1 \pm \sqrt{3}.$$

Donc,

$$\lambda_1 = 1 + \sqrt{3},\ \lambda_2 = 0,\ \lambda_3 = 1 - \sqrt{3},$$

c'est-à-dire que l'on se trouve dans le cas où $\lambda_1 > \lambda_2 > \lambda_3$.

Cherchons le vecteur propre $x^1$. Pour cela formons le système (3):

$$\begin{cases} -\sqrt{3}x_1 + x_2 + x_3 = 0 \\ x_1 - \sqrt{3}x_2 + x_3 = 0 \\ x_1 + x_2 - (1+\sqrt{3})\,x_3 = 0. \end{cases}$$

Les équations de ce système sont deux à deux linéairement indépendantes. En résolvant le système des deux premières équations, on trouve

$$x_1 = \frac{1+\sqrt{3}}{2}\,x_3, \qquad x_2 = \frac{1+\sqrt{3}}{2}\,x_3.$$

Donc, le vecteur

$$y^1 = \left(\frac{1+\sqrt{3}}{2}\,x_3, \quad \frac{1+\sqrt{3}}{2}\,x_3, \; x_3\right)$$

est solution du système. En normant ce vecteur on obtient le vecteur propre

$$x^1 = \frac{y^1}{|y^1|} = \left(\frac{1+\sqrt{3}}{2\sqrt{3+\sqrt{3}}}, \; \frac{1+\sqrt{3}}{2\sqrt{3+\sqrt{3}}}, \; \frac{1}{\sqrt{3+\sqrt{3}}}\right).$$

Cherchons $x^2$ $(\lambda_2 = 0)$:

$$\begin{aligned} x_1 + x_2 + \phantom{0\cdot{}} x_3 &= 0 \\ x_1 + x_2 + \phantom{0\cdot{}} x_3 &= 0 \\ x_1 + x_2 + 0\cdot x_3 &= 0. \end{aligned}$$

En résolvant le système des deux dernières équations (puisque le déterminant des coefficients en $x_2$ et $x_3$ n'est pas nul), on obtient $x_2 = -x_1$, $x_3 = 0$.

Le vecteur $y^2 = (x_1, -x_1, 0)$ est solution du système et le vecteur

$$x^2 = \frac{y^2}{|y^2|} = \left(\frac{1}{\sqrt{2}}, \; -\frac{1}{\sqrt{2}}, \; 0\right)$$

un vecteur unité propre. Il est aisé de voir qu'il est orthogonal à $x^1$ (le produit scalaire de ces vecteurs est nul). Enfin, pour $\lambda_3 = 1 - \sqrt{3}$ on trouve le troisième vecteur propre

$$x^3 = \left(\frac{1}{\sqrt{6+2\sqrt{3}}}, \; \frac{1}{\sqrt{6+2\sqrt{3}}}, \; -\frac{1+\sqrt{3}}{\sqrt{6+2\sqrt{3}}}\right).$$

La matrice orthogonale de passage des coordonnées du vecteur $x = (x_1, x_2, x_3)$ dans le système $(i, j, k)$ aux coordonnées du vec-

teur $x = (\xi_1, \xi_2, \xi_3)$ dans le système $(x^1, x^2, x^3)$ s'écrit

$$\Lambda = \begin{Vmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} \\ x_1^{(3)} & x_2^{(3)} & x_3^{(3)} \end{Vmatrix}$$

$$(|\Lambda| = 1, \ \Lambda^{-1}(i) = x^1, \ \Lambda^{-1}(j) = x^2, \ \Lambda^{-1}(k) = x^3),$$

$$\begin{cases} x_1 = x_1^{(1)}\xi_1 + x_1^{(2)}\xi_2 + x_1^{(3)}\xi_3, \\ x_2 = x_2^{(1)}\xi_1 + x_2^{(2)}\xi_2 + x_2^{(3)}\xi_3, \\ x_3 = x_3^{(1)}\xi_1 + x_3^{(2)}\xi_2 + x_3^{(3)}\xi_3. \end{cases} \tag{6}$$

Cette transformation préserve l'orientation (puisque $|\Lambda| = 1 > 0$). c'est-à-dire que le système $(x^1, x^2, x^3)$ est orienté comme le système initial $(i, j, k)$.

En portant les expressions (6) dans la forme quadratique on obtient la forme canonique

$$(1 + \sqrt{3})\xi_1^2 + (1 - \sqrt{3})\xi_3^2.$$

Etudions maintenant plus en détail les équations 1) à 8) et les surfaces qu'elles décrivent.

L'ellipsoïde

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \qquad (a, b, c > 0). \tag{7}$$

Pour $a = b = c$, l'ellipsoïde (7) se transforme en une sphère de rayon $a$ et de centre $O$, c'est-à-dire en le lieu géométrique des points situés à une distance $a$ de $O$.

Les grandeurs $2a$, $2b$, $2c$ s'appellent *longueurs des axes de l'ellipsoïde*.

L'équation (7) restant invariante par une substitution (simultanée ou séparée) de $-x$ à $x$, de $-y$ à $y$ et de $-z$ à $z$, on conclut que l'ellipsoïde (7) est une surface symétrique par rapport aux plans de coordonnées $x = 0$, $y = 0$ et $z = 0$ et par rapport à l'origine des coordonnées $O$. Il suffit donc d'étudier l'ellipsoïde (7) dans le premier octant, c'est-à-dire pour $x \geqslant 0$, $y \geqslant 0$, $z \geqslant 0$. La portion d'ellipsoïde contenue dans le premier octant est définie par l'équation explicite

$$z = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}, \quad x \geqslant 0, \quad y \geqslant 0, \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1.$$

Pour fixer les idées on admettra que $a \geqslant b \geqslant c$. L'ellipsoïde est une surface bornée. Il est compris dans une boule de centre $O$ et de rayon $a$: les coordonnées $(x, y, z)$ de tout point de l'ellipsoïde vérifient l'inégalité

$$x^2 + y^2 + z^2 \leqslant a^2\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) = a^2 \cdot 1 = a^2.$$

Pour se faire une idée plus exacte de l'ellipsoïde, coupons-le par des plans parallèles aux plans de coordonnées. La section de l'ellipsoïde par les plans $z = h$ $(-c \leqslant h \leqslant c)$ nous donne les ellipses

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1-\frac{h^2}{c^2}$$

d'axes

$$2a\sqrt{1-\frac{h^2}{c^2}}, \qquad 2b\sqrt{1-\frac{h^2}{c^2}}.$$

On remarque que la plus grande ellipse correspond à la section par le plan $z = 0$. On obtient un tableau analogue en coupant l'ellipsoïde par les plans $x = h$ $(-a \leqslant h \leqslant a)$, $y = h$ $(-b \leqslant h \leqslant b)$.

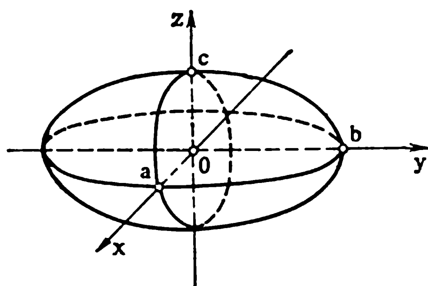



Fig. 151

L'ellipsoïde (7) est de la forme indiquée sur la figure 151.

Les points $(\pm a, 0, 0)$, $(0, \pm b, 0)$, $(0, 0, \pm c)$ sont les *sommets* de l'ellipsoïde.

Si deux axes quelconques sont égaux, l'ellipsoïde (7) est un *ellipsoïde de révolution*, c'est-à-dire la figure engendrée par la rotation d'une ellipse autour du troisième axe.

L'hyperboloïde à une nappe

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad (a, b, c>0). \tag{8}$$

On voit sur cette équation que l'hyperboloïde à une nappe est une surface symétrique par rapport aux plans de coordonnées et par rapport a l'origine des coordonnées. Les nombres $2a$, $2b$, $2c$ sont les *axes de l'hyperboloïde à une nappe*. Les points $(\pm a, 0, 0)$, $(0, \pm b, 0)$ sont les *sommets*.

La section de la surface (8) par le plan $z = h$ est une ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1+\frac{h^2}{c^2}$$

d'axes

$$2a\sqrt{1+\frac{h^2}{c^2}}, \qquad 2b\sqrt{1+\frac{h^2}{c^2}}.$$

Lorsque $h$ varie entre $-\infty$ et $+\infty$, cette ellipse décrit la surface (8).

La section de la surface (8) par un plan $x = h$ (ou $y = h$) est une hyperbole

$$\frac{y^2}{b^2}-\frac{z^2}{c^2}=1-\frac{h^2}{a^2} \quad \left(\frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{h^2}{b^2}\right).$$

Pour $h = \pm a$ la première hyperbole se décompose en deux droites $y = \pm \frac{b}{c} z$.

Si $|h| \leqslant a$, l'axe transverse de l'hyperbole correspondante est une droite parallèle à l'axe $Oy$, pour $|h| > a$, une droite parallèle à l'axe $Oz$ (on appelle *axe transverse* d'une hyperbole l'axe coupé par celle-ci).

Si $a = b$, les sections de la surface (8) par les plans $z = h$ sont des cercles de rayon $a\sqrt{1 + (h^2/c^2)}$. La surface (8) est dans ce cas la surface de révolution de l'hyperbole $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$, $y=0$ autour de l'axe $Oz$. La figure 152 représente un hyperboloïde à une nappe.

L'hyperboloïde à deux nappes

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \qquad (a, b, c > 0). \tag{9}$$

L'équation (9) étant une somme de carrés, la surface qu'elle décrit est symétrique par rapport aux plans $x = 0$, $y = 0$ et $z = 0$ et par rapport à l'origine des coordonnées. L'équation (9) s'écrit encore

$$\frac{y^2}{b^2}+\frac{z^2}{c^2}=-1+\frac{x^2}{a^2}. \tag{9'}$$

On voit que la section de la surface (9′) par un plan $x = h$ ($|h| \geqslant a$) est une ellipse

$$\frac{y^2}{b^2}+\frac{z^2}{c^2}=-1+\frac{h^2}{a^2}$$

d'axes

$$2b\sqrt{(h^2/a^2)-1}, \quad 2c\sqrt{(h^2/a^2)-1}.$$

Pour $|h| < a$, le nombre $(h^2/a^2) - 1$ est strictement négatif, donc la surface (9′) et le plan $x = h$ ne se coupent pas.

La section de la surface (9) par un plan $z = h$ (resp. $y = h$) est une hyperbole

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1+\frac{h^2}{c^2} \quad \left(\text{resp. } \frac{x^2}{a^2}-\frac{z^2}{c^2}=1+\frac{h^2}{b^2}\right).$$

Les points $(\pm a, 0, 0)$ de la surface (9) sont les *sommets de l'hyperboloïde à deux nappes*. La figure 153 représente un hyperboloïde à deux nappes.

Le paraboloïde elliptique

$$\frac{x^2}{p}+\frac{y^2}{q}=2z \quad (p,\ q>0). \tag{10}$$

Les variables $x$ et $y$ figurant par leurs carrés dans (10), cette surface est symétrique par rapport aux plans de coordonnées $x = 0$,

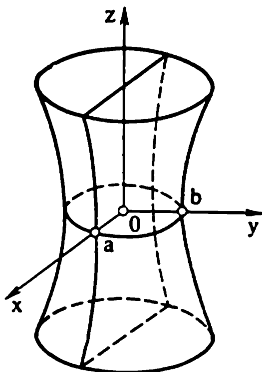



Fig. 152

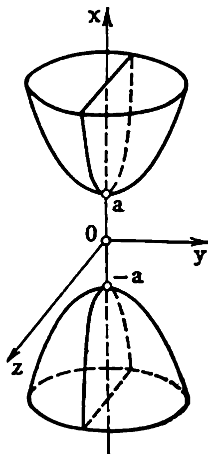



Fig. 153

$y = 0$. Comme $p$, $q > 0$, la surface (10) est située dans le demi-espace $z \geqslant 0$.

La section de la surface (10) par un plan $z = h$ ($h \geqslant 0$) est une ellipse

$$\frac{x^2}{p}+\frac{y^2}{q}=2h$$

d'axes

$$2\sqrt{2ph}, \quad 2\sqrt{2qh}.$$

Lorsque $h$ varie de 0 à $\infty$, cette ellipse décrit la surface (10).

La section de la surface (10) par un plan $x = h$ (resp. $y = h$) est une parabole

$$y^2=2q\left(z-\frac{h^2}{2p}\right) \quad \left(\text{resp. } x^2=2p\left(z-\frac{h^2}{2q}\right)\right)$$

dont le sommet est le point $z=\frac{h^2}{2p}$ $\left(\text{resp. } z=\frac{h^2}{2q}\right)$.

Pour $p = q$, la surface (10) est une surface de révolution engendrée par la rotation de la parabole $x^2 = 2pz$, $y=0$ autour de l'axe $Oz$. On dit que la surface (10) est un *paraboloïde de révolution*.

Le point (0, 0, 0) s'appelle *sommet du paraboloïde elliptique*. La figure 154 représente un paraboloïde elliptique.

Le paraboloïde hyperbolique

$$\frac{x^2}{p}-\frac{y^2}{q}=2z \quad (p, q>0). \tag{11}$$

On voit que la surface définie par (11) est symétrique par rapport aux plans $x=0$ et $y=0$. La section de la surface (11) par un plan $z=h$ est une hyperbole

$$\frac{x^2}{p}-\frac{y^2}{q}=2h.$$

Pour $h>0$, l'axe transverse de l'hyperbole est parallèle à $Ox$,

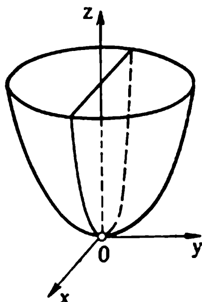



Fig. 154

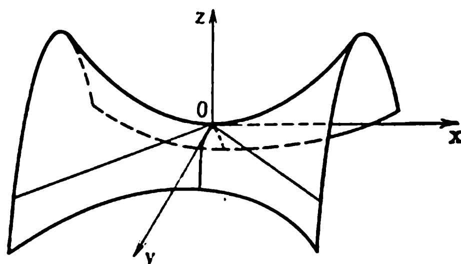



Fig. 155

pour $h<0$, parallèle à l'axe $Oy$. Pour $h=0$ la section est constituée de deux droites concourantes.

La section de la surface (11) par un plan $x=h$ (resp. $y=h$) est une parabole dont la concavité est tournée vers le bas (resp. vers le haut) :

$$-\frac{y^2}{q}=2z-\frac{h^2}{p} \quad \left(\text{resp. } \frac{x^2}{p}=2z+\frac{h^2}{q}\right).$$

La surface (11) est représentée sur la figure 155.

Le cône du second degré

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0 \quad (a, b, c>0). \tag{12}$$

Il est évident que cette surface est symétrique par rapport aux plans $x=0$, $y=0$, $z=0$ et par rapport à l'origine des coordonnées.

La section de la surface (12) par un plan $z=h$ est une ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=\frac{h^2}{c^2}$$

d'axes $2a|h|/c$ et $2b|h|/c$.

La section de la surface (12) par un plan $x = h$ (resp. $y = h$) est une hyperbole

$$\frac{z^2}{c^2} - \frac{y^2}{b^2} = \frac{h^2}{a^2} \quad \left(\text{resp. } \frac{z^2}{c^2} - \frac{x^2}{a^2} = \frac{h^2}{b^2}\right).$$

La section de la surface (12) par un plan $y = hx$ est constituée d'un couple de droites concourantes

$$z = \pm cx\sqrt{(1/a^2) + (h^2/b^2)}.$$

La figure 156 représente un cône.

Le point

$$x^2 + y^2 + z^2 = 0. \tag{13}$$

L'équation (13) n'est vérifiée que par le point $x = y = z = 0$.

Les cylindres du second degré:

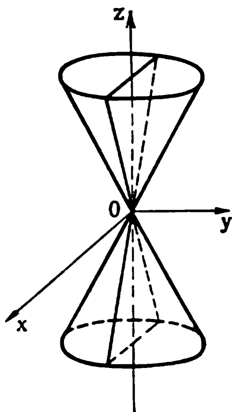



Fig. 156

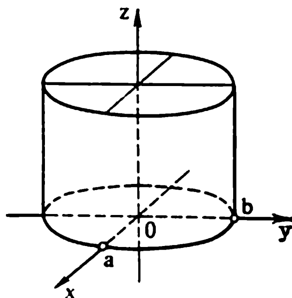



Fig. 157

a) Le cylindre elliptique

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad (a,\ b > 0). \tag{14}$$

L'équation (14) ne contient pas la variable $z$. Elle définit une ellipse d'axes $2a$ et $2b$ sur le plan $xOy$. Si un point $(x, y)$ appartient à l'ellipse, tout point $(x, y, z)$ appartiendra à la surface (14). L'ensemble des points $(x, y, z)$ est une surface engendrée par une droite parallèle à l'axe $Oz$ et s'appuyant sur l'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,\ z = 0. \tag{14'}$$

L'ellipse (14') s'appelle *directrice du cylindre*, la droite, *génératrice du cylindre*.

D'une façon générale, la surface décrite par une droite parallèle à une direction donnée et s'appuyant sur une ligne $L$ est dite *cylindrique*. La surface (14) est représentée sur la figure 157.

b) Les cylindres hyperbolique et parabolique

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \quad (a,\ b>0), \tag{15}$$

$$y^2 = 2px \quad (p>0). \tag{16}$$

Les directrices sont ici une hyperbole et une parabole, les génératrices, des droites parallèles à l'axe $Oz$ et s'appuyant sur les cour-

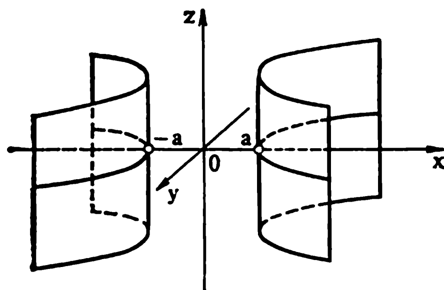


Fig. 158

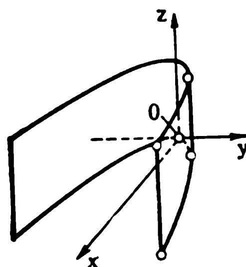


Fig. 159

bes précédentes. Les surfaces (15) et (16) sont représentées sur les figures 158 et 159.

c) Les plans parallèles et concourants. La droite

$$a^2x^2 - b^2y^2 = 0 \quad (a,\ b>0), \tag{17}$$

$$x^2 - a^2 = 0 \quad (a>0), \tag{18}$$

$$z^2 = 0, \tag{19}$$

$$x^2 + y^2 = 0. \tag{20}$$

Les directrices de la surface (17) sont les droites

$$y = \pm\frac{a}{b}x, \quad z=0.$$

Donc, la surface (17) est un couple de plans concourants. Deux variables ne figurent pas dans les équations des surfaces (18) et (19). L'équation (18) définit deux droites $x = \pm a$ dans le plan $xOy$.

Si $x = \pm a$, et $y$ et $z$ sont quelconques, alors les points $(\pm a,\ y,\ z)$ satisfont à l'équation (18), donc la surface (18) est un couple de plans parallèles.

L'équation (19) décrit le plan $xOy$, car elle n'est vérifiée que par les points $(x,\ y,\ 0)$, c'est-à-dire par l'ensemble des points du plan $xOy$.

On peut aussi considérer que $z = 0$ est une directrice dans $xOz$ ou $yOz$ et que les génératrices sont des droites parallèles à $Oy$ ou $Ox$ et s'appuyant sur $z = 0$ dans le plan $xOz$ ou $yOz$.

L'équation (20) est vérifiée par tout point $(0, 0, z)$. Donc c'est l'équation d'une droite, plus exactement de l'axe $Oz$.

S u r f a c e s r é g l é e s.

Certaines quadriques sont engendrées par le mouvement d'une droite, par exemple les surfaces cylindriques et le cône du second degré. Il existe d'autres surfaces engendrées par le mouvement d'une droite.

Une surface engendrée par le mouvement d'une droite s'appelle *surface réglée* et les droites entièrement contenues dans cette surface, *génératrices rectilignes.*

Comme exemples de surfaces réglées on a l'hyperboloïde à une nappe et le paraboloïde hyperbolique.

L'équation de l'hyperboloïde à une nappe (8) s'écrit encore:

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2}$$

ou

$$\left(\frac{x}{a} + \frac{z}{c}\right)\left(\frac{x}{a} - \frac{z}{c}\right) = \left(1 + \frac{y}{b}\right)\left(1 - \frac{y}{b}\right). \qquad (21)$$

Formons le système d'équations du premier degré:

$$\begin{cases} \frac{x}{a} + \frac{z}{c} = k\left(1 + \frac{y}{b}\right), \\ \frac{x}{a} - \frac{z}{c} = \frac{1}{k}\left(1 - \frac{y}{b}\right), \end{cases} \qquad (22)$$

où $k$ est un paramètre.

Le système (22) définit une *famille de droites* de paramètre $k$. Si l'on multiplie membre à membre les équations de (22), on obtient l'équation (21). Donc, tout point $(x, y, z)$ vérifiant le système (22) appartiendra à la surface (21). Donc, chaque droite de la famille (22) est entièrement contenue sur l'hyperboloïde à une nappe.

De façon analogue, le système

$$\begin{cases} \frac{x}{a} + \frac{z}{c} = l\left(1 - \frac{y}{b}\right), \\ \frac{x}{a} - \frac{z}{c} = \frac{1}{l}\left(1 + \frac{y}{b}\right), \end{cases} \qquad (23)$$

où $l$ est un paramètre, définit aussi une famille de droites différente de la famille (22) et située sur la surface (21).

Par chaque point de l'hyperboloïde (21) il passe une droite de chaque famille. Ces deux droites correspondent en général à des valeurs différentes des paramètres $k$ et $l$ (fig. 160). Par exemple, par le point $\left(\sqrt{\frac{3}{2}}\,a,\ \frac{\sqrt{2}}{2}\,b,\ c\right)$ de la surface (21) il passe la droite

de la famille (22) pour $k = (2+\sqrt{6})/(2+\sqrt{2})$ et la droite de la famille (23) pour $l = (2+\sqrt{6})/(2-\sqrt{2})$.

Signalons que les hyperboloïdes à une nappe ont trouvé une application dans le bâtiment. Les tours métalliques qui utilisent les génératrices rectilignes de l'hyperboloïde à une nappe allient la solidité à la sobriété.

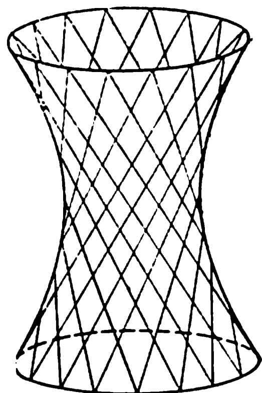


Fig. 160

Il est aisé de vérifier que les deux familles de droites

$$\begin{cases} \dfrac{x}{\sqrt{p}} + \dfrac{y}{\sqrt{q}} = 2kz, \\ \dfrac{x}{\sqrt{p}} - \dfrac{y}{\sqrt{q}} = \dfrac{1}{k}, \end{cases} \qquad (24)$$

$$\begin{cases} \dfrac{x}{\sqrt{p}} + \dfrac{y}{\sqrt{q}} = \dfrac{1}{l}, \\ \dfrac{x}{\sqrt{p}} - \dfrac{y}{\sqrt{q}} = 2lz \end{cases} \qquad (25)$$

forment un paraboloïde hyperbolique (11).

Les droites des familles (24) et (25) sont situées sur cette surface et inversement tout point de cette surface est l'intersection d'une droite de (24) et d'une droite de (25).

## § 26. Théorie générale des quadriques

Soit donnée une quadrique

$$\sum_{k=1}^{3} \sum_{l=1}^{3} a_{kl} x_k x_l + 2 \sum_{l=1}^{3} A_l x_l + B = 0, \qquad (1)$$

où $a_{kl} = a_{lk}$, $A_l$ et $B$ sont des constantes.

Au § 25 on a énuméré 8 cas particuliers de l'équation (1) et on a mentionné que l'équation (1) sous réserve qu'elle ne soit pas l'équation d'une surface imaginaire se ramenait à l'un de ces cas particuliers dans un certain système de coordonnées rectangulaires.

On démontre ce fait plus bas.

On commence par étudier la forme quadratique du premier membre de l'équation (1).

Le théorème 2 du § 22 nous dit que cette forme se réduit par la transformation orthogonale

$$x_i = \sum_{s=1}^{3} \beta_{is} x'_s \qquad (i = 1, 2, 3) \qquad (2)$$

à la forme:

$$\sum_{k=1}^{3}\sum_{l=1}^{3} a_{kl}x_k x_l = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2,$$

où $\lambda_1$, $\lambda_2$, $\lambda_3$ sont des nombres réels connus.

A noter que les formules (2) sont les formules de passage du système de coordonnées rectangulaires $(x_1, x_2, x_3)$ à un autre système de coordonnées rectangulaires $(x_1', x_2', x_3')$.

Dans les nouvelles coordonnées. la surface a visiblement pour équation

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2 + 2\sum_{k=1}^{3} A_k' x_k' + B = 0, \tag{3}$$

où $A_k'$ sont des constantes.

Traitons le cas où les trois nombres $\lambda_1$, $\lambda_2$, $\lambda_3$ sont différents de 0 ($\lambda_i \neq 0$, $i = 1, 2, 3$).

Plaçons l'origine du nouveau système de coordonnées $(\xi_1, \xi_2, \xi_3)$ au point $(a_1, a_2, a_3)$. On a

$$x_i' = a_i + \xi_i \qquad (i = 1, 2, 3).$$

L'équation de la surface s'écrit dans les nouvelles coordonnées:

$$\lambda_1 (a_1 + \xi_1)^2 + \lambda_2 (a_2 + \xi_2)^2 + \lambda_3 (a_3 + \xi_3)^2 + 2\sum_{k=1}^{3} A_k' (a_k + \xi_k) + B = 0$$

ou bien

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 + 2\sum_{l=1}^{3} (a_l \lambda_l + A_l')\, \xi_l + B_1 = 0,$$

où $B_1$ est une constante. Si l'on pose

$$a_l = -\frac{A_l'}{\lambda_l} \qquad (l = 1, 2, 3),$$

cette équation prend la forme simple:

$$\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \lambda_3 \xi_3^2 = -B_1. \tag{4}$$

Supposons que les nombres $\lambda_1$, $\lambda_2$, $\lambda_3$ sont de même signe.

Si $B_1 = 0$. alors l'équation (4) est vérifiée par *un seul point*, le point (0, 0, 0) (voir 7), § 25).

Si $B_1 \neq 0$ et est du signe de $\lambda_1$, $\lambda_2$, $\lambda_3$, il est évident qu'il n'existe pas de points de coordonnées réelles vérifiant l'équation (4). La *surface* (1) est *imaginaire.*

Si $B_1 \neq 0$ et n'est pas du signe de $\lambda_1$, $\lambda_2$, $\lambda_3$, alors l'équation (4) peut se mettre sous la forme

$$\frac{\xi_1^2}{-\frac{B_1}{\lambda_1}}+\frac{\xi_2^2}{-\frac{B_1}{\lambda_2}}+\frac{\xi_3^2}{-\frac{B_1}{\lambda_3}}=1; \tag{5}$$

si l'on pose

$$a^2=-\frac{B_1}{\lambda_1}, \qquad b^2=-\frac{B_2}{\lambda_2}, \qquad c^2=-\frac{B_3}{\lambda_3},$$

elle devient

$$\frac{\xi_1^2}{a^2}+\frac{\xi_2^2}{b^2}+\frac{\xi_3^2}{c^2}=1 \qquad (a,\ b,\ c>0).$$

Donc, la surface (1) est un *ellipsoïde* (voir 1), § 25).

Supposons maintenant que deux des nombres $\lambda_1$, $\lambda_2$, $\lambda_3$ sont de même signe.

Si $B_1 = 0$, on peut admettre dans l'équation (4) que $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 < 0$, quitte à multiplier (4) par $-1$ et quitte à permuter les $\xi_j$. En posant

$$\lambda_1=\frac{1}{a^2}, \quad \lambda_2=\frac{1}{b^2}, \quad \lambda_3=-\frac{1}{c^2} \quad (a,\ b,\ c>0),$$

on obtient l'équation du *cône* (voir 6), § 25)

$$\frac{\xi_1^2}{a^2}+\frac{\xi_2^2}{b^2}-\frac{\xi_3^2}{c^2}=0.$$

Pour $B_1 \neq 0$ on se sert de nouveau de la formule (5). Distinguons deux cas foncièrement différents:

$$\frac{\xi_1^2}{a^2}+\frac{\xi_2^2}{b^2}-\frac{\xi_3^2}{c^2}=1 \quad \text{(\textit{l'hyperboloïde à une nappe}, voir 2), § 25)},$$

$$\frac{\xi_1^2}{a^2}-\frac{\xi_2^2}{b^2}-\frac{\xi_3^2}{c^2}=1 \quad \text{(\textit{l'hyperboloïde à deux nappes}, voir 3), § 25)}.$$

Les autres cas se ramènent à ces deux derniers par un changement de coordonnées $\xi_i$ convenable.

Supposons maintenant que $\Delta = \lambda_1\lambda_2\lambda_3 = 0$. L'un au moins de ces nombres, pour fixer les idées $\lambda_3$, est nul.

Traitons d'abord le cas où $A_3' = 0$. L'équation (3) s'écrit

$$\lambda_1 x_1'^2 + \lambda_2 x_2'^2 + 2\,(A_1' x_1' + A_2' x_2') + B = 0,$$

où les nombres $\lambda_1$, $\lambda_2$, $A_1'$, $A_2'$, $B$ sont arbitraires.

Dans le plan $(x_1', x_2')$ c'est l'équation d'une conique. Dans l'espace rapporté aux coordonnées $(x_1', x_2', x_3')$ c'est l'*équation d'une surface cylindrique* (voir 8), § 25) dont la directrice est une *conique* et la génératrice une droite parallèle à l'axe des $x_3'$ (voir 8), § 25).

On admettra dans la suite que $A'_3 \neq 0$ et $\lambda_3 = 0$. L'équation (3) devient alors

$$(\lambda_1 x_1'^2 + 2A'_1 x'_1) + (\lambda_2 x_2'^2 + 2A'_2 x'_2) + 2A'_3 x'_3 + B = 0. \qquad (3')$$

On distinguera les cas suivants, les autres s'y rapportant par un changement de coordonnées ou par une multiplication par $-1$: a) $\lambda_1, \lambda_2 > 0$; b) $\lambda_1 > 0$, $\lambda_2 < 0$; c) $\lambda_1 > 0$, $\lambda_2 = 0$.

a) $\lambda_1, \lambda_2 > 0$. Mettons $\lambda_1$ et $\lambda_2$ en facteur et complétons les expressions entre parenthèses à des carrés parfaits. Comme $\lambda_1, \lambda_2$ et $A'_3$ ne sont pas nuls, on a

$$\lambda_1 (x'_1 + \alpha)^2 + \lambda_2 (x'_2 + \beta)^2 + 2A'_3 (x'_3 + \gamma) = 0,$$

où, pour simplifier l'écriture, on a désigné par $\alpha$, $\beta$ et $\gamma$ les expressions obtenues. En posant

$$\xi = x'_1 + \alpha, \quad \eta = x'_2 + \beta, \quad \zeta = x'_3 + \gamma,$$

on obtient

$$\lambda_1 \xi^2 + \lambda_2 \eta^2 = -2A'_3 \zeta$$

ou

$$\frac{\xi^2}{-\frac{A'_3}{\lambda_1}} + \frac{\eta^2}{-\frac{A'_3}{\lambda_2}} = 2\zeta. \qquad (6)$$

Si $-A'_3 > 0$, l'équation (6) est de la forme

$$\frac{\xi^2}{p} + \frac{\eta^2}{q} = 2\zeta \text{ (\textit{le paraboloïde elliptique}, voir 4), § 25)} \qquad (7)$$

$$(p, q > 0).$$

Si $-A'_3 < 0$, en remplaçant $\zeta$ par $-\zeta$, on obtient de nouveau une équation de la forme (7), c'est-à-dire un *paraboloïde elliptique*.

b) $\lambda_1 > 0$, $\lambda_2 < 0$ $(A'_3 \neq 0)$. On se sert de l'équation (6). Si $A'_3 < 0$, cette équation s'écrit

$$\frac{\xi^2}{p} - \frac{\eta^2}{q} = 2\zeta \text{ (\textit{le paraboloïde hyperbolique}, voir 5), § 25)} \qquad (8)$$

$$(p, q > 0).$$

Si $A'_3 > 0$, en permutant $\xi$ et $\eta$ on obtient une équation de la forme (8), c'est-à-dire un *paraboloïde hyperbolique*.

c) $\lambda_1 > 0$, $\lambda_2 = 0$ $(A'_3 \neq 0)$. L'équation (3′) devient

$$(\lambda_1 x_1'^2 + 2A'_1 x'_1) + 2 (A'_2 x'_2 + A'_3 x'_3) + B = 0.$$

En procédant comme pour cas a) on ramène cette équation à la forme

$$\lambda_1 (x'_1 + \alpha)^2 + 2A'_2 x'_2 + 2A'_3 (x'_3 + \beta) = 0,$$

où $\alpha$ et $\beta$ sont des nombres. En posant

$$\xi = x'_1 + \alpha, \quad \eta = x'_2, \quad \zeta = x'_3 + \beta,$$

on met cette équation sous la forme

$$\lambda_1\xi^2 + 2(A_2'\eta + A_3'\zeta) = 0. \qquad (9)$$

Dans le plan rapporté aux coordonnées $(\eta, \zeta)$ considérons le vecteur $\omega = (A_2', A_3')$. Ecrivons-le sous la forme

$$(A_2', A_3') = \rho(\cos\alpha, \sin\alpha),$$

où $\rho > 0$ est le module de $\omega$, $(\cos\alpha, \sin\alpha)$ le vecteur unité de $\omega$. Le vecteur $(\sin\alpha, -\cos\alpha)$ est aussi un vecteur unité qui est orthogonal au premier.

Soit dans le plan rapporté au système de coordonnées rectangulaires $(\eta, \zeta)$ la transformation orthogonale

$$u = \eta\cos\alpha + \zeta\sin\alpha, \qquad v = \eta\sin\alpha - \zeta\cos\alpha.$$

Cette transformation transforme le système de coordonnées rectangulaires $(\eta, \zeta)$ en un système de coordonnées rectangulaires $(u, v)$, et le système de coordonnées rectangulaires $(\xi, \eta, \zeta)$ en le système de coordonnées rectangulaires $(\xi, u, v)$. En définitive, l'équation (9) que l'on peut écrire

$$\lambda_1\xi^2 + 2\rho(\eta\cos\alpha + \zeta\sin\alpha) = 0$$

prend la forme suivante:

$$\lambda_1\xi^2 + 2\rho u = 0,$$

ou en remplaçant $u$ par $-u$:

$$\xi^2 = 2pu \quad (p = \rho/\lambda_1 > 0),$$

ou enfin en permutant $\xi$ et $u$:

$$u^2 = 2p\xi \quad (p > 0),$$

on reconnaît l'*équation d'un cylindre parabolique* (dans les coordonnées rectangulaires $(\xi, u, v)$).

Nous avons étudié tous les cas particuliers de l'équation (1) et pour chacun d'eux nous avons trouvé un système de coordonnées rectangulaires dans lequel l'équation (1) prend l'une des formes 1) à 8) du § 25. Ce que nous voulions.

# INDEX

**SOMMAIRE DU TOME II**

ÉQUATIONS DIFFÉRENTIELLES ORDINAIRES
INTÉGRALES MULTIPLES
ANALYSE VECTORIELLE
SÉRIES DE FOURIER ET INTÉGRALE DE FOURIER
ÉQUATIONS DE PHYSIQUE MATHÉMATIQUE
THÉORIE DES FONCTIONS D'UNE VARIABLE COMPLEXE
CALCUL OPÉRATIONNEL

## À NOS LECTEURS

Les Editions Mir vous seraient très reconnaissantes de bien vouloir leur communiquer votre opinion sur le contenu de ce livre, sa traduction et sa présentation, ainsi que toute autre suggestion.

*Notre adresse*:
2, Pervi Rijski péréoulok,
Moscou, I-110, GSP, U.R.S.S.

*Imprimé en Union Soviétique*